

Mekanistisen selittämisen haasteita ja kehityslinjoja

*Onttisuus, mekanistinen mallintaminen ja mutuaalinen manipuloitavuus uudessa
mekanistisessa filosofiassa*

Jan-Joel Karvinen
Helsingin yliopisto
Humanistinen tiedekunta
Teoreettinen filosofia
Pro gradu -tutkielma
Helmikuu 2019

Sisällys

1. Johdanto	1
2. Mekanistisen filosofian historialliset ja käsitteelliset taustat	2
2.1. Peittävän lain malli, lyhyt katsaus	2
2.1.1. Peittävän lain mallin kestättömyys	5
2.2. Selityksen kausaalinen luonne	8
2.3. Kohti mekanistista selittämistä	9
2.3.1. Onttiset ja episteemiset sudenkuopat	10
2.3.2. Konstitutionaalinen ja kausaalinen selittäminen	13
3. Mekanismit – Se mikä selittää	16
3.1. Tapahtuman alaiset entiteetit	16
3.2. Järjestytyneisyys: avaruudellisuus, ajallisuus ja hierarkkisuus	21
4. Mallit – Se millä selitetään	26
4.1. Mitä mallit ovat, lyhyt yleinen kuvaus	26
4.2. Mekanistiset mallit ja niiden ulottuvuudet	27
4.3. Mekanistisen mallintamisen ja ontisuuden kritiikki	32
4.4. Miksi mallintamisen tulee olla ontista?	35
5. Manipuloitavuus - Se mikä osoittaa riippuvaisuuden	38
5.1. Kausaaliset rakenneyhtälöt	38
5.2. Interventionismi: kausaalisuuden kontrafaktuaalinen analyysi	39
5.3. Mikä tekee interventionismista houkuttelevan?	45
6. Interventionismin soveltaminen konstitutionaaliseen analyysiin	48
6.1. Mutuaalinen manipulaatio	48
6.2. Mutuaalisen manipulaation dilemma	55
6.2.1. Paksukätinen ratkaisu	58
6.2.2. Kausaalinen ratkaisu	63
6.3. Mutaalisen manipulaation tulevaisuudennäkymät	67
7. Päätelmiä	68
8. Lähdeluettelo	70

1. Johdanto

Tämän tarkastelun keskiössä ovat kolme tieteenfilosofista haastetta. Niitä ei voida pitää tieteenfilosofian ainoina haasteina. Väitän kuitenkin, että ne ovat oleellisesti jokaisen tieteellisen selittämisteorian haasteita. Nämä kolme haastetta voidaan muotoilla seuraavien kolmen kysymyksen muodossa:

H1) Mikä selittää maailman olemisen niin-tai-näin?

H2) Millä keinoilla ja missä muodossa selitys voidaan ja tulisi muotoilla?

H3) Millä perusteella annetun selitysmuotoilun sisältö on selityksellinen?

Kutsun ensimmäistä *ontologiseksi (onttiseksi) haasteeksi*, toiseksi mainittua *epistেমiseksi (mallintamisen) haasteeksi* ja viimeiseksi mainittua *relevanssin haasteeksi*. Keskeinen väitteeni on, että näihin kolmeen kysymykseen voidaan vastata mekanistisella selittämisen teoriolla. Mekanistinen filosofia on kirjava joukko erinäisiä näkemyksiä, eikä mekanistien keskuudessa näytä olevan yksimielisyyttä siitä, kuinka yllä esitettyihin haasteisiin voitaisiin vastata. Tarkasteluni teoreettisena pohjana tulee pitkälti olemaan Carl Craverin (2006; 2007; 2014) filosofia. Tarkoitukseni ei ole omaksua tätä teoreettista pohjaa ilman kritiikkiä, sillä tulen muodostamaan synteessin Craverin näkemyksien sekä sitä kohtaan esitettyjen kriittisten näkemyksien pohjalta.

Tarkasteluni tulee ensimmäisen haasteen suhteen nojautumaan siihen Wesley Salmonin (1984) esittämään onttiseen kantaan, jonka monet niin kutsutut ”uudet mekanistit” ovat omaksuneet (etenkin Craver). Kannan mukaan mekanismit, jotka tullaan esittämään maailman kausaalisten (eli ontittiset) rakenteiden kokonaisuuksina, selittävät maailman olemisen niin-tai-näin. Määrittelen tarkemmin ontittisuuden osiossa 2, jossa myös tuon ilmi, että kausaalinen (etiologinen) ja konstitutionaalinen selittäminen ovat kaksi ontittisen selittämisen muotoa. Vastaavasti osiossa 3 esitän mekanismien määritelmän.

Väitän, että ontittisen kannan hyväksyminen määrittää sitä, kuinka haasteisiin H2 ja H3 voidaan vastata. Sanalla sanoen osiossa 4 tullaan väittämään, että mekanistiset tai kausaaliset selitykset voidaan hahmottaa mekanismeja representoivilla mekanistisilla malleilla. Tulen väittämään, että käsitykset ei-mekanistisista selitysmalleista, kuten minimalistista tai puhtaasti dynaamisista malleista, ovat ongelmallisia sikäli kuin niissä selitysvoinan alkuperä joudutaan mystifioimaan. Perustan tämän väitteelle, että

selityksellisen mallin tarkoitus on ensisijassa jäljittää *kausaalisia* riippuvaisuussuhteita puhtaasti *matemaattisten* riippuvaisuussuhteiden sijaan.

Osiot 5 ja 6 tulevat keskittymään viimeiseen H3-haasteeseen. Osiossa 5 tuon ilmi, kuinka monet uudet mekanistit ovat väittäneet, omaksuessaan James Woodwardin (2001; 2002; 2003) interventionistisen käsityksen kausaalisuudesta, että mekanistien selityksen relevanssiehto voidaan muotoilla kontrafaktuaalisten riippuvaisuussuhteiden avulla. Tulen esittämään, että mallin sisältämät muuttujasuhteet ovat selityksellisesti relevantteja, mikäli ne jäljittävät kausaalisia riippuvaisuussuhteita, jotka mahdollistavat vastaamisen ”mitä-jos-maailma-olisi-toisenlainen”-kysymyksiin. Osiossa 6 tulen tarkastelemaan, kuinka Craver on pyrkinyt soveltamaan tätä ajatusta myös konstitutionaaliseen selittämiseen niin sanotulla ”mutuaalisella manipulaatio”-ehdolla. Vetoan kuitenkin tuoreessa kirjallisuudessa esiintyneisiin kantoihin (Leuridan 2012; Romero 2015; Baumgartner & Gebharder 2017), joiden nojalla esitän Craverin MM-ehdon olevan kestävä. Tähän kestävämyyteen tarjoan kaksi tuoreessa kirjallisuudessa esiintynyttä ratkaisuehdotusta, joita voidaan kutsua abduktiiviseksi (Baumgartner & Casini 2017) ja kausaaliseksi ratkaisuksi (Krickel 2017; 2018). Tulen kuitenkin väittämään, että nämä ehdotukset eivät ole täysin tyydyttäviä.

Pyrin arviomaan vastauksia esitettyihin haasteisiin naturalistisen tieteenfilosofian perspektiivistä käsin. Perspektiiville ominaista on tarkastella tiedettä niin kuin se näyttäytyy tieteellisen käytännön valossa, ei niin kuin se näyttäytyy filosofin analyttisissä pohdinnoissa. Tämä näkyy myös pyrkimyksessäni hakea ”filosofinen evidenssi” esitetyille kannoille erityistieteisiin (lähinnä biologiaan) nojautuvista yksinkertaistetuista esimerkeistä. Tarkastelu on kuitenkin syytä aloittaa tarkastelemalla mekanistisen filosofian lähtökohtia eli klassista peittävän lain mallia ja sen kritiikkiä. Mekanistisen filosofian voidaan katsoa olevan tästä kritiikistä noussut kanta. Seuraavaksi esitän peittävän lain malli lyhyesti Carl Hempelin ja Paul Oppenheimien näkemyksien mukaisesti.

2. Mekanistisen filosofian historialliset ja käsitteelliset taustat

2.1. Peittävän lain malli, lyhyt katsaus

Perinteisen käsityksen mukaan selitys voidaan jakaa kahteen peruselementtiin ja nämä ovat *explanandum* (selitettävä) ja *explanans* (selittävä). Peittävän lain mallin mukaan ensimmäinen on lause, joka viittaa selitettävään ilmiöön, kun taas jälkimmäinen on se

lauseiden joukko, josta ensin mainittu lause johdetaan loogisesti (Hempel & Oppenheim 1948, 137). Käytännössä peittävän lain mallin mukaan jokainen selitys on näin ollen looginen argumentti. Yhtäältä se voi olla deduktiivisnomologinen (*deductive nomological*; jatkossa DN) ja toisaalta tilastollinen/induktiivinen argumentti (*inductive statistical*; jatkossa IS). Molemmat argumenttityypit ovat peittävän lain mallin mukaan selityksiä, joissa ilmiö johdetaan annetun loogisen periaatteen mukaisesti premisseistä.

Explanansin lausejoukko voidaan jakaa edelleen kahteen alajoukkoon; yhtäältä joukkoon, joka sisältää lauseet (C1, C2 ... Cn), jotka ilmaisevat selitettävän ilmiön taustaehdot, lisäksi *explanansiin* sisältyy myös joukko, joka sisältää lauseet (L1, L2 ... Ln), jotka edustavat lakiväitteitä (ibid.). Lisäksi on välttämätöntä muotoilla seuraavat loogiset ehdot *explanandum*in ja *explanansin* välisestä suhteesta (vrt. Hempel & Oppenheim 1948, 137):

P1: *Explanandum* on *explanansin* looginen seuraus

P2: *Explanansin* tulee sisältää lakiväitteet, joista *explanandum* voidaan johtaa loogisesti.

P3: *Explanansin* tulee olla sisällöltään empiirinen.

P4: Lauseet, jotka konstituoivat *explanansia*, ovat tosia.

Näiden ehtojen nojalla voidaan muotoilla seuraava DN-malli skeema (ibid., 138):

$$\frac{L1, L2, \dots, Ln}{C1, C2, \dots, Cn} \\ E$$

Mikäli *explanandum* on annettu mutta *explanansin* taustaehdot eivät ole tiedetty, silloin voidaan puhua ilmiön ”takautuvasta” selityksestä ja mikäli *explanandumia* edeltävät lauseet ovat annettu mutta *explanandum* ei ole tiedetty, silloin voidaan puhua varsinaisesta selityksestä (ibid.). Tästä voidaan huomata, että käytännössä peittävän lain mallissa ennustus ja selitys ovat loogisesti isomorfisia sikäli kuin mallin mukaan ne ovat loogiselta rakenteeltaan identtisiä. Carl Hempel toteaa (1965, 366-7), että DN-selitys ja ennustus eroavat käsitteinä oikeastaan vain ”pragmaattisilta” aspekteiltaan.

Sekä takautuvien että tyypillisten ennustusten voidaan molempien katsoa olevan DN-mallin tärkeitä ominaisuuksia. DN-mallin mukaan tässä ei ole mitään mystistä, koska loppujen lopuksi kyseessä on vain selityksien johtaminen loogisesti empiirisistä

havainnoista. Selitykset ovat näin ollen empiriasta loogisesti derivoitavissa olevia vastauksia ”miksi”-kysymyksiin (Hempel 1965, 334, 412). Voidaan esimerkiksi kysyä, miksi X tapahtui. Tällöin DN-mallin mukaisesti selityksen muoto on takautuva: siitä, että asianlaidan tiedetään olevan X, seuraa loogisesti se, että asianlaidan on täytynyt olla se-ja-se ennen X:n ilmenemistä. Vastaavasti voidaan kysyä, millä perusteella Y tulee tapahtumaan. Tällöin DN-mallin mukaisesti selityksen muoto on ennustuksellinen; siitä, että asialaidan tiedetään olevan nyt niin-tai-näin, seuraa loogisesti se, että asianlaita tulee olemaan Y. Tämä looginen seuraavuus on puolestaan DN-mallissa jotakin, minkä luonnonlait takaavat, ja sitä voidaankin kutsua *noomiseksi odotettavuudeksi* (ks. Salmon 1989, 48). Selitys on siis odotettavuuden osoittamista; jotakin tapahtui, koska se oli odotettua sen-ja-sen nojalla.

Näinmuodoin DN-mallia voidaan luonnehtia seuraavalla symmetriateesillä (Hempel 1965, 367; ks. myös Salmon 1989, 48): i) jokainen pätevä DN-selitys on potentiaalinen ennustus ja ii) jokainen ei-tilastollinen ennustus on DN-selitys. Mainittu ”ei-tilastollinen ennustus” antaa implisiittisesti ymmärtää, että on olemassa selityksiä, jotka ovat potentiaalisia *tilastollisia* ennustuksia. Tällöin voitaisiin puhua selityksestä, joka ei ole DN-muotoa vaan sen sijaan edellä mainittua IS-muotoa.¹

IS-selityksissä suhde, joka vallitsee *explanansin* ja *explanandum*in välillä, ei ole deduktiivinen vaan sen sijaan induktiivinen. Tämä ero voidaan todeta myös niin, että deduktiivisuutta voidaan luonnehtia *totuuden* säilyttävänä relaationa, kun taas induktiivisuutta voidaan luonnehtia *korkean todennäköisyyden* säilyttävänä relaationa. Tällöin voidaan sanoa, että IS-selityksessä *explanans* antaa (*confers*) korkean todennäköisyyden *explanandumille* (Hempel 1965, 379, 384). Sama voidaan ilmaista myös seuraavalla skeemalla (Hempel 1965, 390):

$$p(G|F) = r \\ \frac{F(a)}{G(a)} [r]$$

Skeemassa ensimmäinen premissi ilmaisee tilastollisen yleistyksen, jossa ilmaistaan *G:n* todennäköisyys *r*, kun *F* on annettu. IS-selityksen kannalta keskeistä on se, että *r*

¹ Hempel erottaa ns. deduktiivistilastolliset (DS) selitykset IS-selityksistä (1965, 380-1). Ensin mainitussa tilastollinen ennustus johdetaan deduktiivisesti tilastollisista laeista ja taustaehdoista. Mainitun käsitteleminen ei ole kuitenkaan aiheen kannalta oleellista sikäli kuin DS-selityksien muoto on sama kuin DN-mallien.

ilmaisee korkean todennäköisyyden. Tosin peittävän lain mallin kannattajat eivät näyttäneet koskaan päässeen yksimielisyyteen siitä, miten r tulisi eksplikoida eli, mikä on IS-selityksen kannalta *tarpeeksi* korkea todennäköisyys. Toinen premissi ilmaisee puolestaan sen, että asianlaita on a :n suhteen F , josta voidaan johtaa (induktiivisesti) lähes varmuudella, että asianlaita on a suhteen myös G . Toisin sanoen jakoviivalla ilmoitettu $[r]$ ilmaisee sen, että ilmoitettu päätelmä: ”kun asianlaita on a :n suhteen F , silloin asianlaita on myös a :n suhteen G ”, on johdettu r -todennäköisyyden nojalla.

Yhteenvedona DN- ja IS-selitysmallien luonnehdinnasta voidaan johtaa niitä yhdistävä ominaisuus, jonka peittävän lain malli esittää tieteellisen selityksen ominaisuutena:

D1: Tieteellinen selitys on argumentti, joka osoittaa ilmiön noomisen tai tilastollisen odotettavuuden.

Toisin sanoen D1 tuo ilmi sen, että peittävän lain mallissa *explanansin* ja *explanandumin* välillä on sellainen looginen suhde, joka muistuttaa loogisesti ennustuksen mahdollistavien tekijöiden ja ennustettavan (ilmiön) tekijän välistä loogista suhdetta. Seuraavaksi on tarkoitus siirtyä tarkastelemaan, mitä kestäättömiä johtopäätöksiä annetusta D1:stä voidaan johtaa.

2.1.1. Peittävän lain mallin kestäättömyys

Peittävän lain malli on ollut monenlaisen kritiikin kohteena, eikä tässä voida toivoa kaiken kattavaa esitystä sen puutteellisuuksista. Tarkoituksena on sen sijaan tuoda ilmi muutamia puutteellisuuksia, jotka ovat mekanistisen filosofian kannalta keskeisiä. Nämä puutteellisuudet koskevat lähinnä peittävän lain-mallin kyvyttömyyttä erottaa ennustuksellisuus ja selittäminen toisistaan. Seuraavaksi esitettävä ensimmäinen vastaesimerkki (Scriven 1959, 480) on loppujen lopuksi seurausta juuri siitä, että peittävän lain-mallissa vain ennustettavuudella on selitysarvoa.

(2.1.2.1): *Korrelaatio*. Kun luotettavasti toimivan barometrin elohopea laskee nopeasti, voidaan tästä päätellä se, että myrsky on tuloillaan. Barometrin lukeman laskun ja myrskyn ilmenemisen välillä vallitsee siis suhde, joka on ilmaistavissa universaalisti kvantifioitun proposition avulla: ”aina kun luotettavan barometrin lukema laskee, silloin myrsky on tuloillaan”. Peittävän lain mallissa tämä propositio voidaan mieltää lakiväitteeksi, josta voidaan johtaa myrskyn selitys. Johtopäätös on kuitenkin kestäättön, sillä ei ole mielekäästä väittää, että barometrin elohopean lasku sinänsä

selittää myrskyn. Sen sijaan mielekkäämpää on sanoa, että se, mikä *aiheuttaa* barometrin lukeman laskun ja myrskyn, selittää sekä myrskyn että elohopean laskun. Sanalla sanoen elohopean lasku ja myrsky ovat yhden ja saman seikan (ilmanpaineen muutoksen) kaksi erillistä seurausta.

Barometrin lukema mahdollistaa myrskyn ennustamisen, mutta se ei mahdollista myrskyn selittämistä, koska se ei ole myrskyn aiheuttava tekijä; kummallakin mainituista on nimittäin yhteinen *selittävä* syy eli ilmanpaineen muutos. Korrelaatiolla on ennustuksellista voimaa mutta ei selitysvoimaa, tosin todetun perusteella on ilmeistä, että peittävän lain mallissa ei voida ottaa siis huomioon sitä, etteivät korrelaatit *selitä* toinen toistaan (Scriven 1959, 480). Ongelmallisuus on suoraa seurausta siitä, että D1:n mukaisesti peittävän lain mallin mukaan selitysvoima on loppujen lopuksi odotettavuutta.

Peittävän lain malli on myös ongelmallinen sen takia, että se ei ole yhteensopiva sellaisten skenaarioiden kanssa, joissa ilmiö voidaan selittää, mutta jossa sen ilmeneminen ei ole todennäköistä (Scriven 1959, 480). Tämä käy ilmi, kun tarkastellaan ilmiöitä, joiden ilmeneminen on ”luotettavasti” epätodennäköistä. Esimerkkinä voidaan käyttää välittäjäaineiden epätodennäköistä vapautumista hermosolusta (Bogen 2005, 400; Craver 2007, 26).

(2.1.2.2): *Epätodennäköisyys*. Välittäjäaineiden vapautumisessa on kyseessä neurobiologinen ilmiö, joka on hyvin karkeasti sanoen seurausta hermosolun toimintapotentiaalista, joka on puolestaan hermosolun depolarisoitumisesta seuraava sähkökemiallinen impulssi. Vaikka toimintapotentiaali selittää välittäjäaineiden vapautumisen, ei se tee siitä todennäköistä. Itseasiassa välittäjäaineiden vapautumisen todennäköisyys vaihtelee hyvin suuresti ja tämä vaihtelu on seurausta useista eri tekijöistä, mutta suurimmassa osassa keskushermostoa impulssin saapuminen aksoniterminaaliin ei johda välittäjäaineiden *todennäköiseen* vapautumiseen. Vapautumisen selittävä tekijä (mekanismi) on tässä tapauksessa siis epäluotettava.

Peittävän lain (IS) mallin mukaan toimintapotentiaali ei olisi välittäjäaineiden vapautumisen selittävä tekijä, koska sen perusteella välittäjäaineiden vapautuminen ei ole todennäköistä tai odotettua. Tämän kaltainen johtopäätös on kuitenkin nykyisen tieteen valossa kestävä. Pikemminkin se, mitä halutaan sanoa, on se, että toimintapotentiaali voi olla selityksellinen, koska se on jossakin määrin ”eron tekijä”

välittäjäaineiden vapautumisen suhteen. Eron tekevyys (*difference making*)² ei kuitenkaan edellytä, että eron tekijän nojalla ilmiö ilmenee suurella todennäköisyydellä. Näin ollen on todettu, että ilmiön ilmenemisen korkealla todennäköisyydellä sinänsä ei ole selitysvuimakkuutta lisäävää vaikutusta (Salmon 1984, 32; 1989, 59–60).

Eron tekevyyydestä päästään kolmanteen ongelmaan. Voidaan nimittäin osoittaa peittävän lain mallia koskeva ongelma, joka koskee selityksellistä relevanssia. Kysymys on siitä, mitkä asiat ovat selityksen kannalta relevantteja ja jotka niin ollen tulisi ottaa huomioon ilmiön selityksessä. Ilmeistähän on se, että selityksessä halutaan vedota vain ilmiön kannalta relevantteihin seikkoihin. Peittävän lain mallissa ei ole kuitenkaan mitään, minkä nojalla tämän kaltainen relevanssiehto voitaisiin ilmaista (Salmon 1984 31; 1989, 50). Tämä käy ilmeiseksi seuraavasta vastaesimerkistä

(2.1.2.3.): *Irrelevanssi*. Seuraava vasta-argumentti voidaan muotoilla niin että (vrt. Salmon 1971, 34):

[L1]: <i>Kaikki miehet, jotka syövät säännöllisesti ehkäisypillereitä, eivät tule raskaaksi</i>
[C1]: <i>X on mies</i>
[C2]: <i>X syö säännöllisesti ehkäisypillereitä</i>
[E]: <i>X ei tule raskaaksi</i>

Mainittu ”selitys” on pätevä argumentti, mutta selityksellisesti se on omituinen, koska tässä tapauksessa ehkäisypillereiden syöminen on selityksen kannalta irrelevanttia. Sanalla sanoen ehkäisypillereiden syöminen ei lisää selitettävään ilmiöön mitään relevanttia, jolloin sen sijoittaminen tai sijoittamatta jättäminen *explanansiin* on yhdentekevää. Peittävän lain mallin ongelma puolestaan syntyy siitä, ettei sillä ole sellaista käsitteellistä aspektia, jonka nojalla edellä mainittu ilmeinen ero relevanttien ja irrelevanttien seikkojen välillä voitaisiin ilmaista eksplisiittisesti.

Viimeiseksi voidaan mainita se ongelma, että peittävän lain mallissa selitys on aina tosi deduktiivinen argumentti, joka ei sinällään ota huomioon selityksen suuntaa.

Kuuluisana esimerkkinä tästä on ”Brombergerin lipputanko” -vastaesimerkki (Salmon 1989, 47; vrt. Bromberger 1966), joka voidaan esittää seuraavasti.³

² Tälle termille ei näytä olevan vakiintunutta suomennosta, joten olen ottanut ”vapauden” tämän *eron tekijä* -termin käytön suhteen. Tuleva tarkastelu lieene osoittavan valinnan kelvollisuuden.

³ Lieene aiheellista tuoda ilmi, ettei ko. vastaesimerkki ole varsinaisesti Brombergerin esittämä, toisin kuin usein kirjallisuudessa esitetään. Myös Bromberger itse on tuonut ilmi tämän väärinkäsityksen varsin eksplisiittisesti (kts. 1992). Hän kuitenkin toteaa esittäneensä samankaltaisen vasta-argumentin.

(2.1.2.4): *Symmetria*. Oletetaan, että lipputangon pituus on X ja oletetaan lisäksi, että aurinko paistaa taivaalla niin, että sen johdosta valontulokulma suhteessa lipputankoon on θ :stä. Käytännössä tämä tarkoittaa sitä, että lipputangon varjon pituus Y , voidaan johtaa deduktiivisesti X :stä ja θ :stä. Peittävän lain mallin mukaan tämä on validi argumentti, jonka johtopäätös voidaan johtaa (ennustaa) deduktiivisesti sen premisseistä, mikä tarkoittaa sitä, että kyseessä on peittävän lain mallin mukainen selitys. Toisin peittävän lain mallilla on esimerkin kannalta omituisia seurauksia; malli mahdollistaa käänteisen skenaarion eli sen, että X johdetaan Y :stä ja θ :stä. Sanalla sanoen malli mahdollistaa lipputangon pituuden ”selittämisen” sen lipputangon varjolla. Tämä on kuitenkin kestävä johtopäätös sikäli kuin on ilmeistä, että X ja θ selittävät Y :n mutta ei toisin päin. Tämän vastaesimerkin pohjalta voidaan tehdä se päätelmä, että selitys on asymmetrinen suhde kahden seikan välillä: jos A voi selittää B :n, tarkoittaa se sitä, että B ei voi selittää A :ta.

2.2. Selityksen kausaalinen luonne

Edellä esitetty peittävän lain mallin kritiikki osoittaa selvästi muutamia mielenkiintoisia tieteellisen selittämisen ominaisuuksia, jotka ovat jääneet peittävän lain mallissa vaille niiden ansaitsemaa huomiota. Nämä voidaan karkeasti listata seuraavasti:

- S1: Saman syyn erilliset seuraukset eivät selitä toinen toisiaan. [2.1.2.1]
- S2: Selityksen ei tarvitse osoittaa, että *explanans* on *explanandum*in nojalla todennäköinen. [2.1.2.2]
- S3: Vain relevantit seikat ovat selitysvomaisia. [2.1.2.3]
- S4: Selityksen suunta on sama kuin kausaliteetin suunta [2.1.2.4]

Yksinkertaisesti mainittujen ominaisuuksien voidaan katsoa osoittavan sen, että tieteellisellä selittämällä on jotakin tekemistä kausaalisuuden kanssa; selitys identifioi ne syyt ja kausaaliset suhteet, jotka vastaavat selitettävästä ilmiöstä (Salmon 1989, 46). Näyttäisi siltä, että symmetria-argumentista voidaan myös päätellä, ettei ns. *loogisen empirismin kolmas dogma* ole kestävä selitys ole vain looginen argumentti (Salmon 1998, 102-3). Vasta-argumentit osoittavat sen, että selityksessä halutaan jäljittää eron tekijät eli käytännössä osoittaa mikä riippuu mistäkin ja selittää seuraukset niiden syillä, jotka tuottavat nämä eron tekevät ssuhteet (vrt. Salmon 1989, 45). Lisäksi ne osoittavat oleellisesti sen, että ilmiön noominen odotettavuus ei sinänsä lisää mitään selitysvomaa, eikä se ole sen vuoksi riittävä tai välttämätön

ehto selitykselle. Selitysvoima ei näinmuodoin nouse siitä, että voidaan osoittaa asioiden tapahtuvan säännöllisesti.

(S1)-(S4) tuovat ilmi sen useille selityksille oleellisen ominaisuuden, että tieteellisen selityksen tulisi osoittaa, mikä aiheuttaa (kausaalisesti) ilmiön ja mikä on sen kannalta eron tekevää (Salmon 1989, 45; 1998; 309). David Lewisin mukaan kausaalisella eron tekevyydellä tarkoitetaan tarkemmin sanoen seuraavaa:

“We think of a cause as something that makes a difference, and the difference it makes must be a difference from what would have happened without it.” (Lewis 1986, 167).

Peittävän lain mallin kyvyttömyys ottaa huomioon selityksen kausaalinen luonne, johtaa siihen, että kanta on osittain liian löyhä tieteellisen selittämisen suhteen. Pelkkä odotettavuus ei ole riittävä eikä edes välttämätön vaatimus. Lienee siis tuskin liioiteltua väittää, että tieteellisessä selittämisessä ollaan itseasiassa kiinnostuneita siitä, mikä maailman kausaalisissa rakenteissa mahdollistaa tämän ennustettavuuden. Sanalla sanoen usein tieteellinen mielenkiinto kohdistuu ilmiöiden konstitutionaalisiin osiin. Tarkastellaan seuraavaksi, miten todettu voidaan ymmärtää mekanistisen selittämisen teoriassa.

2.3. Kohti mekanistista selittämistä

Selityksellinen malli on sellainen, että siinä todenmukaisesti identifioidaan tai osoitetaan ne maailman rakenteet, jotka vastaavat selitettävästä ilmiöstä, ja tätä voidaan kutsua *onttiseksi selittämisenäkemykseksi* (Salmon 1984, 262-3; Craver 2007, 200; 2014, 39). Aiheen kannalta keskeinen väite on se, että tämän osoittamisen kohteena ovat ilmiöstä vastaavat *kausaaliset mekanismit* ja niiden toiminnalliset osat. Suurin osa niin kutsutuista ”uusista mekanisteista” näyttäisivät omaksuneet tämän kannan.⁴ Uusien mekanistien keskeinen argumentti onkin se, että käytännössä erityistieteiden, kuten esimerkiksi biologian alalla, puhe laeista on viimekädessä aina puhetta mekanismeista ja niiden toiminnallisista osista (Thagard 1998, 73; Machamer *et al.* 2000, 4; Glennan 2002, 348; Craver 2007, 7-8; Bechtel 2008, 10). Tässä mielessä DN-mallin

⁴ Termiä ”uusi mekanismi” käytetään usein nykyisessä mekanistisessa kirjallisuudessa viittaamaan lähinnä mekanistiseen kirjallisuuteen, jota on julkaistu 1990-luvulta lähtien. On vaikeaa antaa yhdistävää kuvausta tälle kirjallisuudelle, mutta usein ominaista tälle kirjallisuudelle on sen keskittyminen erityistieteiden mekanistiseen analyysiin ja niiden tieteelliseen praktiikkaan. (Tarkempi analyysi termin alkuperästä ks. Glennan 2017, §1.2. ja §1.3).

lakiväitevaatimus on siis itseasiassa liian vahva, jos halutaan sallia erityistieteelliset selitykset. Useat filosofit ovat nimittäin tuoneet esille sen, että peittävät luonnonlait ovat varsin harvassa erityistieteellisillä aloilla (Woodward 2001; vrt. Cartwright 1983, 45). On myös väitetty, että parhaimmillaan yleistykset voidaan muotoilla erityistieteissä, niin että ne koskevat vain rajattuja systeemejä, so. ne eivät päde yhtä laajojen olosuhdemuutoksien suhteen kuin luonnonlait (Mitchel 1997; 2000). Tässä vaiheessa oleellista on kuitenkin tehdä selväksi se, mitä tarkoitetaan useiden mekanistien kannattamalla ontillisella selittämiskannalla.

2.3.1. *Onttiset ja episteemiset sudenkuopat*

Onttinen selittämiskanta on näkemys, jossa selitys ei ole argumentti, vaan eräänlainen ”osoitus” (*exhibition*) (Salmon 1984, 296). Mikäli mekanismit ymmärretään selityksen kantaviksi entiteeteiksi, voidaan tämän kaltaista selittämistä kutsua alkuperäiseksi Salmonin muotoilemaksi onttiseksi (ontic) kannaksi. Kenties parhaiten Salmon muotoilee kantansa kuvatessaan Alberto Coffan (1974) filosofiaa onttisena kantana:

“The linguistic entities that are often called ‘explanations’ are statements reporting the actual explanation. Explanations [...] are fully objective and, where explanations of nonhuman facts are concerned, they exist whether or not anyone ever discovers or describes them” (Salmon 1989, 133).

Käytännössä kyse on siitä, että *explanandum* paikannetaan maailman kausaaliseen rakenteeseen, jolloin onttisessa selittämisen kanssa ilmiön selittävä tekijä on jokin maailmassa oleva entiteetti (riippuvuussuhde). Selitys ei ole siis looginen argumentti, niin kuin peittävän lain mallissa. Sen sijaan selitys ymmärretään mielestä riippumattomana, objektiivisena riippuvaisuusrelaationa eli eron tekävänä suhteena asioiden välillä. Toisin sanoen tämän kaltainen, kenties naiivi onttilaisen selittämisen kanta on eräänlainen episteemisen selittämisen anti-teesi. Onttilaisen tulkinnan on katsottu kuitenkin johtavan kategoriseen virhepäätelmään (Betchel & Abrahamsen 2005, 424; kts. Wright 2012). Tulen kuitenkin osoittamaan, että tämä tulkinta koskee vain ja ainoastaan naiivia tulkintaa onttilaisuudesta. Paremminkin sanoen käsitys, jonka mukaan selitys on mielestä riippumaton, ei-representationaalinen relaatio, on kestävä johtopäätös. Tämä johtuu yksinkertaisesti siitä, että tieteellinen selitys ja mekanismi eivät ole päällekkäisiä (Wright 2012, 383). Päällekkäisyydellä tarkoitetaan tässä sitä, että tieteellisellä selityksellä on yleisesti katsottu olevan ominaisuuksia, joita

mekanismeilla, mielestä riippumattomina entiteetteinä ei ole. Esimerkiksi välittäjäainetransmissiomekanismin kausaalisella selityksellä on *episteemisiä ominaisuuksia* (totuudenmukaisuus, oikeutuksen aste jne.), joita ei voida sanoa olevan välittäjäainetransmissiomekanismilla itsellään.

Onttisen kannan kestättömyyden seurauksena on katsottu olevan se, *reductio ad absurdum*, että selityksien täytyisi periaatteessa edeltää itse mekanismin tieteellistä havaintoa (Betchel & Abrahamsen 2005, 425). Näin ollen Huntingtonin taudin mekanistisen selityksen voitaisiin katsoa olevan olemassa jopa sellaisessa mahdollisessa skenaariossa, jossa Huntingtonin taudin tuottavaa geenivirhettä ei olisi koskaan löydetty. Oleellista tämän vasta-argumentin mukaan on se, että mekanismi on saattanut olla olemassa kauan ennen sen tieteellistä löytämistä, mutta tästä ei voida johtaa sitä, että asianlaita olisi samoin ilmiön selityksen suhteen. Sama pätee kääntäen: mekanismi saattaa lakata olemasta, mutta tämä ei tarkoita sitä, että sen joskus mahdollistama selitys lakkaisi olemasta sen myötä (ibid.). Todettu näyttäisi johtavan episteemiseen kantaa, jossa tieteellinen selitys on mielestä riippuvainen entiteetti (Wright 2012, 376). Tämän ei kuitenkaan tarvitse tarkoittaa sitä, että on välttämätöntä siirtyä takaisin DN-mallin inferentiaaliseen kantaan, jossa selitys on argumentti. Voidaan sanoa, että selitys ratsastaa mekanismia sinänsä representoivan mekanistisen mallin mukana (Bechtel 2008, 18; Betchel & Abrahamsen 2005, 425). Representoimalla ilmiön mahdollistamaa mekanismia, mallin tarkoituksena on mahdollisimman todenmukaisesti osoittaa maailman kausaalinen rakenne.

Toisaalta ontittisuuden ”neo-salmonilaiset” kannattajat eivät näytä hyväksyvän johtopäätöstä, jonka mukaan selitys on vain episteeminen entiteetti (Craver 2007, 27-8; 2014). Neo-salmonilaisten mukaan puhdas episteemisen selitysmalli johtaa omiin sudenkuoppiinsa. Väitteen mukaan, mikäli selitys ymmärretään mielestä riippuvaiseksi, on seuraavia väittämiä on mahdotonta ymmärtää mielekkäästi: a) on olemassa havaitsemattomia ilmiöitä, joille on olemassa selitys; b) on olemassa ilmiöitä, joita nyt ei voida selittää, mutta joille on olemassa selitys; c) tieteen päämääränä on *löytää* selitys ilmiöille, ei siis vain muotoilla selitystä; d) on olemassa mahdollisesti niin kompleksisia ilmiöitä, joita ei voida koskaan selittää, mutta joille on olemassa ainakin periaatteessa selitys (Craver 2014, 36). Toisin sanoen ontittisen kannan ja väittämien (a)-(d) hylkääminen johtaisi siihen, että tieteelliselle selitykselle ei olisi selkeää objektiivista pohjaa. Oleellista on huomata, että selityksellisiä tekstejä tai representaatioita

arvioidaan aina suhteessa siihen, missä määrin ne ovat todenmukaisia (ibid., 37). Tosin jos ontisuus hylätään täysin, katoaa silloin se pohja, jota vasten mainittua todenmukaisuutta voidaan arvioida. Sanalla sanoen ontisuus näyttäisi mahdollistavan pohjan sellaiselle normatiiviselle kriteerille, jonka kautta voidaan arvioida sitä, missä määrin annettu tieteellinen selitys kertoo jostakin maailman seikasta.

Mekanistinen filosofia jakautuu siis kahteen leiriin sen perusteella, miten ne suhtautuvat ontiseen kantaan. Yhtäältä on mekanisteja, jotka hyväksyvät Salmonin ontisen tulkinnan (Craver 2007, 200; 2014; Glennan 2002, 343; Woodward 2003, 202; Machamer et al. 2000, §3). Toisaalla on taas mekanistien vähemmistö, joka eksplisiittisesti hylkää tämän kannan ja korvaa sen episteemisellä selittämisen kannalla (Bechtel & Abrahamsen 2005; Bechtel 2008; Wright 2012). Toisaalta on myös väitetty, että keskustelu ontisen-episteemisen-dikotomian ympärillä vaikuttaa olevan vailla mielekästä pohjaa, sillä mikäli kumpaakaan selityskantaa ei lueta kirjaimellisesti, ovat ne loppujen lopuksi toisiaan täydentäviä (Glennan 2017, 219-223; Illari 2013). Toisin sanoen erimielisyydet syntyvät siitä, että ontisen ja episteemisen katsotaan olevan kaksi toisensa poissulkevaa seikkaa, vaikka loppujen lopuksi ne ovat vain kaksi toisiaan täydentävää ulottuvuutta. Tämä täydentävyys voidaan ilmaista seuraavien ehtojen valossa (vrt. Illari 2013, 250):

Onttinen ehto: selityksen täytyy osoittaa maailman kausaalinen rakenne niiden (mekanististen) konstitutionaalisten osien kautta, jotka vastaavat selitettävästä ilmiöstä.

Episteeminen ehto: selityksen täytyy muodostaa malli eli representaatio siitä, kuinka ilmiö on osiensa konstituoina niin, että ilmiö on sen nojalla manipuloitavissa, ymmärrettävissä ja kommunikoitavissa tiedeyhteisössä.

Episteeminen ulottuvuus tuo ilmi sen, että selitys tarvitsee aina välikappaleen eli representaation, jonka nojalla jokin toimija välittää ymmärrettävän selityksen muille toimijoille. Onttinen ulottuvuus täydentää tätä tuomalla ilmi intersubjektiiviset kriteerit selityksen arvioimiselle. Loppujen lopuksi kiista ontisten ja episteemisten kantojen välillä näyttäisi olevan vain verbaalinen. Kyse on siitä kumpaa ulottuvuutta (onttista tai episteemistä) painotetaan. Mikäli nämä kaksi ulottuvuutta voidaan ymmärtää edellä kuvatun tapaan, niin ettei toista ulottuvuutta painoteta ylitse toisen, voidaan välttyä yhdeltä filosofiselta keskustelulta, joka ei varsinaisesti tuo mitään lisää itse

mekanistisen filosofian keskusteluun. Aiheen kannalta on kuitenkin vielä keskeistä erotella ne kaksi selityksellistä osoittamisen tapaa, jotka ovat molemmat ontittisia. Kyse on nimittäin kausaalisten konstitutionaalisten selityksien erosta.

2.3.2. Konstitutionaalinen ja kausaalinen selittäminen

Tarkastellaan, miten konstitutionaalinen ja kausaalinen selittäminen eroavat toisistaan. Molemmissa selitystavoissa selitys kumpuaa maailman kausaalisen rakenteen osoittamisesta (Salmon 1984, 270). Konstitutionaalisen selittämisen keskiössä on kuitenkin ilmiön käyttäytyminen sen konstitutionaalisten osien nojalla; kausaalinen (etiologinen) selittäminen osoittaa puolestaan ilmiön kausaalisen historian (ibid.). Mainittua eroa näiden kahden selittämisen välillä voidaan myös ilmentää niin, että siinä missä kausaaliset selitykset selittävät ilmiön osoittamalla (*exhibiting*) sen kausaalisen historian, jonka nojalla ilmiö on tullut tuotetuksi, selittävät konstitutionaaliset selitykset ilmiön puolestaan osoittamalla sen kausaalisen verkoston, joka vastaa ilmiöstä (Craver 2007, 107-108). Konstitutionaalinen selitys on siis luonteeltaan sellainen, että se osoittaa sen, minkä takia jollakin asioilla on niille ominaiset toiminnalliset taipumukset (dispositiot) (Ylikoski 2013, 279; Cummins 1983)⁵, kun taas kausaalinen selitys voidaan tulkita niin, että se osoittaa, mitä tästä toiminallisuudesta seuraa tai voi seurata.

Oleellista on siis huomata, että nämä kaksi selitysmallia jäljittävät kahta eri riippuvaisuussuhdetta (Ylikoski 2013, 289). Kausaalinen selitys jäljittää asioiden välistä kausaalista riippuvaisuutta, kun taas konstitutionaalinen jäljittää tasojen välistä (järjestyksellistä) riippuvaisuutta. Kausaalisen ja konstitutionaalisen selittämisen erosta voidaan antaa seuraava filosofinen esimerkki (Ylikoski 2013, 279): etiologisesti voidaan selittää, miksi lasi rikkoutui tai miksi siitä on tullut rikkoutuvan hauras, mutta tämä selitys ei tee selkoa siitä, mikä tekee lasista rikkoutuvan hauraan (mahdollistaa lasin disposition tai taipumuksen hajota tietyn ärsykkeen vaikutuksesta). Voidaan siis erottaa kolme mielekästä miksi-kysymystä (ibid.):

Wh-1: Miksi lasista on tullut hauras?

Wh-2: Miksi lasi rikkoutui?

Wh-3: Mikä tekee lasista hauraan tai rikkoutuvan?

⁵ Tämä dispositionaalinen tulkinta ei ole kuitenkaan yksimielisesti hyväksytty mekanistien keskuudessa. Esimerkiksi Marie Kaiser ja Beate Krickel (2017) antavat kattavan esityksen eri tulkinnoista ja niiden metafysisistä haasteellisuuksista.

Kausaalinen selitys vastaa kysymyksiin 1 ja 2, mutta se ei voi vastata kysymykseen 3. Käytännössä kausaalinen selittäminen siis vastaa vain siihen, mikä on lasin rikkoutumisen suhteen kausaalisesti relevanttia. Konstitutionaalinen selitys puolestaan vastaa kysymykseen, miksi lasilla on rikkoutumisen mahdollistava mahdollistama kausaalinen kapasiteetti. Toisin sanoen sen päämääränä on tuoda ilmi rikkoutumisen mahdollistava molekyylylityyppien joukko, joka tietyllä tavalla järjestäytyneenä tekee lasista hauraan.

Wesley Salmonin mukaan (1984, 270) kattava tieteellinen selitys on sellainen, että se jäljittää sekä ilmiön kausaalisen että konstitutionaalisen aspektin. Nämä kaksi aspektia ovat loppujen lopuksi toinen toisiaan täydentäviä. Tosin uusi mekanistinen kirjallisuus näyttäisi merkittävältä osin keskittyvä konstitutionaaliseen aspektiin (Darden 2006; Craver 2007; Bechtel 2008; Glennan 2017). Tämä johtunee siitä, että konstitutionaalinen selittäminen tuo esille sen tärkeän selittämisen aspektin, joka on keskeinen suurimmalle osalle erityistieteen aloista, kuten esimerkiksi biologialle.⁶ Toki yhtäältä tieteellinen mielenkiinto voi kohdistua siihen kausaaliseen aspektiin, minkä takia jokin partikulaarinen biologinen X-systeemi ilmentää kausaalista kapasiteettiaan tavalla Y. Toisaalta tieteellinen mielenkiinto voi ja usein näyttäisikin kohdistuvan siihen, mikä mahdollistaa sen, että X-tyyppiset biologiset systeemit voivat ilmentää kausaalista kapasiteettiaan Y (vrt. Ylikoski 2013, 280). Katsoisin, että viimeiseksi mainittu aspekti on ensin mainittua syvällisempi, minkä väitän käyvän selväksi, kun tarkastellaan seuraavia kysymyksiä:

Wh-4: Miksi X:llä on valkuaisaineita virtsassa (proteinuria)?

Wh-5: Mikä rooli munuaisten hiussuonikeräsillä on verenpuhdistusmekanismin kannalta?

Jos verrataan esitettyjä kysymyksiin edellä esitettyihin (Wh-1-3), niin huomataan että Wh-4 on jokseenkin yhtäläinen Wh-2 kanssa ja Wh-5 puolestaan Wh-3 kanssa.

Oleellista on huomata, että molemmat mainituista kysymyksistä ovat tieteellisesti mielenkiintoisia – ensimmäinen patologisesti, kun taas jälkimmäinen enemmänkin anatomisesti. Tieteenfilosofisesti nämä kysymykset ovat mielenkiintoisia sikäli kuin

⁶ Fundamentaalista fysiikkaa voidaan kenties pitää poikkeuksellisenä erityistieteenä sikäli kuin sen alalla näyttäisi olevan määritelmällisesti ongelmallista lähestyä ilmiöitä konstitutionaalisesti. Toisin sanoen jos tutkittava ilmiö on jokin todellisuuden fundamentaalinen rakenne, silloin ei ole enää käsitteellisesti mielekäästä kysyä: ”mikä konstituoit tutkittavaa ilmiötä?” Jos kysymys on mielekäs, silloin ”fundamentaali rakenne” ei ole kovin fundamentaalinen.

niiden muoto valottaa sitä, mihin tieteellinen mielenkiinto kohdistuu. Voidaan esimerkiksi väittää, että vastausta Wh-4 kohdan kysymykseen on vaikeaa ymmärtää sen kaikessa syvyydessään, mikäli Wh-5 kohdan kysymykseen ei ole jo tarjottu vastausta. Toisin sanoen X:llä on todennäköisesti valkuaisaineita virtsassaan, koska X:llä on nefropatia eli X:n munuaisten hiussuonikeräset ovat vaurioituneet. Tämä puolestaan tarkoittaa sitä, että X:n verenpuhdistusmekanismi ei toimi normaalisti, mikä on valkuaisaineen esiintymisen virtsassa *kausaalinen selitys*. Selityksen sisältöä on kuitenkin mahdotonta ymmärtää tieteellisellä syvällisyydellä, mikäli ei ymmärretä sitä, miten verenpuhdistusmekanismin tulisi normaalisti toimia tai mikä *konstituo*i normaalia verenpuhdistusmekanismeja. Käytännössä kyse on siitä, mikä rooli hiussuonikeräsillä on mainitun mekanismin kannalta. Sanalla sanoen verenpuhdistusmekanismin kausaalinen rakenne on se, minkä nojalla Wh-4 kuvaama partikulaarinen tapahtuma on mahdollinen. Lisäksi mikäli halutaan suorittaa lääketieteellinen interventio nefropatian hoitamiseksi, tulee vastauksen Wh-5 kysymykseen olla jo selvillä. Tieteellisesti *kontrolloitu interventio*⁷ ei ole mahdollinen ennen kuin tiedetään, mitkä toiminnalliset osat konstituoivat normaalia verenpuhdistusmekanismeja.

Todetun perusteella ehdotan seuraavaa tieteellisen selityksen ominaisuutta edellä mainittujen (S1-S4) lisäksi:

S5: Perusteellinen (mekanistinen) selitys jäljittää ilmiöstä vastaavan mekanismin konstitutionaaliset osat.

Esitetty (S5) tuo ilmi sen, mihin uusi mekanistinen filosofia pyrkii tieteellisen selityksen analyysissä. Nyt, kun ollaan tarkennettu se, mitä tarkoitetaan ontillisella, kausaalisella ja konstitutionaalisella selittämällä, voidaan syventyä itse mekanistiseen analyysiin. Mekanistisen selittämisen voidaan katsoa sisältävän kaikki mainitut kolme tieteellisen selittämisen aspektia. Oleellista on kuitenkin huomata, että usein selityksen kausaalinen tai etiologinen aspekti jää uudessa mekanistisessa analyysissä sivuun. Seuraavaksi on tarkoitus paneutua siihen, minkälainen muoto mekanismeilla on, minkä jälkeen voidaan palata analysoimaan tarkemmin sitä, mitä tarkoitetaan konstitutionaalisuudella.

⁷ Tässä kontrolloidulla interventiotermillä halutaan tehdä selväksi se, että kyseessä on tieteelliseen ymmärrykseen perustuva interventio (ymmärretään sen mekanismin toiminnallisuus, jota ollaan manipuloimassa). Tämä voidaan erottaa interventiotyypeistä, jossa mekanismi ei ole ymmärretty; esimerkiksi tapauksesta käy alkoholin tai useiden muiden päihteiden historiallinen käyttö keskushermoston manipuloimiseksi kauan ennen kuin ymmärrettiin ne mekanismit, jotka mahdollistavat nämä manipulaatiosuhteet.

3. Mekanismit – Se mikä selittää

3.1. Tapahtuman alaiset entiteetit

Lienee vain johtuvan arkipäiväisen ajatuksen intuitiivisuudesta, että mekanismien usein ajatellaan olevan koneiden tapaisia kokonaisuuksia. Tämä intuitiivisuus juontaa juurensa varmaankin siitä, että useimmat asiat, joita arjessa nimitetään mekanismeiksi, ovat juurikin koneita. Tosin mekanismit, joita tässä tullaan analysoimaan, eivät ole aina koneita.

Biologit, puhuessaan esimerkiksi ihmisestä biologisena mekanismina, eivät näin puhuessaan puhu ihmisestä vieterikellon tapaisena koneena. Valaiseva huomiona koneiden ja mekanismien erosta voisi olla kenties se, että jopa kone, joka lakkaa toimimasta, lakkaa samalla olemasta mekanismi, vaikkakin sitä voidaan yhä nimittää (rikkinäiseksi) koneeksi (vrt. Darden 2006, 280-1). Sanalla sanoen mekanismi on laajempi käsite kuin koneen käsite. Seuraavaksi onkin tarkoitus tehdä selväksi se, kuinka laajasta käsitteestä on kysymys. Tarkastelu lienee aiheellista aloittaa niistä muutamista määritelmistä, joita kirjallisuudessa on tarjottu. Seuraavien luonnehdintojen voidaan katsoa olevan uuden mekanismin keskeisimpiä:

“Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions.” (Machamer et al. 2000, 3.)

”A mechanism for a behavior is a complex system that produces that behavior by the interaction of a number of parts, where the interactions between parts can be characterized by direct, invariant, change-relating generalizations.” (Glennan 2002, 344.)

“A mechanism is a structure performing a function in virtue of its component parts, component operations, and their organization. The orchestrated functioning of the mechanism is responsible for one or more phenomena.” (Bechtel & Abrahamsen 2005, 423.)

Tarkastellaan ensin, mikä näitä kantoja erottaa, sillä mekanistit eivät ole yksimielisiä mekanismin käsitteestä. Yhtäältä sen on sanottu olevan *systeemi*, *rakenne* tai *input-output järjestelmä*, mutta nämä kaikki kolme esitettyä luonnehdintaa ovat ongelmallisia.

Käsitykset siitä, että mekanistinen kokonaisuus olisi systeemi (Glennan 2002) tai rakenne (Bechtel & Abrahamsen 2005), ovat ongelmallisia sikäli kuin niiden perusteella

mekanismit olisivat joustamattomia kokonaisuuksia. Useat mekanismit ovat nimittäin dynaamisia ja joustavia kokonaisuuksia sikäli kuin mekanismit saattavat tuottaa tuotteita, jotka ovat sinällään sen osia (Illari & Williamson 2012, 121). Esimerkiksi proteiinisynteesimekanismissa jotkin mekanismin relevanteista entiteeteistä tuotetaan niin sanotusti ”lennossa”; näin on laita muun muassa lähetti-RNA:n suhteen, joka useimmissa tapauksissa degeneroituu, kun se on suorittanut tehtävänsä proteiinisynteesissä (Darden 2006, 281). Se, mitä tästä voidaan päätellä, on se, että monet biologiset mekanismit tuottavat niiden omat ”rakenteensa”, jotka usein myös tuhoutuvat jossakin mekanismin vaiheessa.

Lisäksi käsitys siitä, että mekanismi olisi eräänlainen alku- ja lopputilan välikappale (Machamer et al. 2000; Machamer 2004), on ongelmallinen käsitys sikäli kuin kaikkia mekanismeja ei voida kuvata etenevän lineaarisesti input-ärsykkeistä ja output-tuotteisiin. Toisin sanoen mainittu käsitys ei ota huomioon, että esimerkiksi useat metaboliset mekanismit ovat *kehäprosesseja* (positiiviset ja negatiiviset *feedback*-loopit) (Illari & Williamson 2012, 122; Bechtel 2011). Esimerkiksi sitruunahapposykli on kehäprosessille perustuva mekanismi, joka ei varsinaisesti etene input-lähtökohdasta erilliseen output-tuotteeseen. Sitruunahapposyklimekanismissa sitruunahappo on nimittäin useiden kemiallistenreaktioiden kohde, joista suurin osa on hapettumisreaktioita; ”lopputuloksena” on oksaloetikkahappo, joka yhdessä asetyylikoentsyymi-A kanssa muodostaa uuden sitruunahappomolekyylin, joka on seuraavan vastaavan kaltaisen syklin ”lähtökohta” (Bechtel 2011, 545; 2008, 213-4).

Samoin käsitys, jonka mukaan mekanismin toiminta on säännöllinen (Machamer et al. 2000; Machamer 2004), on ongelmallinen (Bogen 2005). Käsitykseen voidaan soveltaa edellä esitettyä välittäjäainetransmissio esimerkkiä. Kyse on siitä, että myöskään kausaalinen selitys ei edellytä luonnon säännönmukaisuutta. Vaikka luonto usein toimii säännönmukaisesti, ja tämä helpottaa tieteen tekemistä, ei siitä seuraa, että syy- ja seuraussuhteet ovat välttämättä säännönmukaisia. Jim Bogen on tiivistänyt tämän kritiikin hyvin todetessaan: ”It’s just that causality is one thing, and regularity, another.” (2005, 399). Annetun kritiikin perusteella kestävämpänä ratkaisuna on esitetty, että mekanismit *vastaavat* ilmiöstä (Illari & Williamson 2012, 120, 124). Tämä määritelmä on siitä oivallinen, että se ei ole yhtä poissulkeva ja rajoittunut kuin edellä mainitut määritelmät, jolloin mekanismeiksi voidaan lukea myös sellaiset tapaukset, joissa

jotakin ilmiötä *suoritetaan, tuotetaan, säädellään, kontrolloidaan ja ylläpidetään*. Tästä päästään siihen, että mekanismit ovat hyvin monimuotoinen joukko⁸.

Kun siirrytään tarkastelemaan sitä, mistä mekanismit koostuvat, voidaan todeta edeltävien lainausten perusteella, että mekanistit ovat varsin yksimielisesti dualisteja. Vaikka yllä olevissa lainauksissa on eroja sen suhteen, miten mekanismit ilmenevät, näyttäisivät mekanistit olevan jokseenkin yhtä mieltä siitä, että karkeasti sanoen jokainen mekanismin aito komponentti koostuu kahdesta erillisiin ontologisiin kategorioihin kuuluvasta seikasta:

Komponentti-def₁: jokainen mekanismin komponentti koostuu: a) muutoksen kohteena olevasta seikasta ja b) muutokseen aikaan saavasta seikasta.

Tätä ”entiteetti-aktiiviteetti”-dualismia voidaan kutsua *perusdualismiksi*. Huomioitavana poikkeuksena tähän yksimielisyyteen voidaan pitää kuitenkin James Woodwardin (2002; 2011) esittämää mekanismin määritelmää, jossa dualistinen keskustelu ohitetaan. Mainitussa määritelmässä mekanismien komponentit määritellään kontrafaktuaalisen analyysin mukaisesti invarianssin ja interventioiden avulla. Tässä ei voida mennä syvemmin siihen, kuinka kestävä Woodwardin ”ontologisesti sitoutumaton” mekanismin määritelmä on; mutta ilman kritiikkiä sekään ei ole selvinnyt (Waskan 2011; ks. myös Glennan 2017, §6.7)

Yhtä kaikki vaikka mekanistit ovat jokseenkin yksimielisiä perusdualismista, mekanistit eivät näytä olevan kovinkaan yksimielisiä siitä, mitä mainittu dualinen kokonaisuus pitää sisällään. Perusdualismin valossa oleellista on huomata, että sen mukaan mekanismin komponentti on muodoltaan sellainen, että siinä entiteetit ovat muutoksen kohde – tai entiteetti on jonkin toiminnan alainen. ”Aktiiviteettien” on siis ehdotettu olevan muutoksen mahdollistavia tekijöitä (Machamer et al. 2000, 3). Dualismin mukaan mekanismit ovat siis aina aktiivisia sikäli kuin *mekanismit tekevät asioita* (ibid., 5). Tällöin esimerkiksi pysähtynyt kello ei ole mekanismi sinänsä, koska se ei ole aktiivinen entiteetti (Darden 2006, 280). Keskeisen tästä huomiosta tekee siis se, etteivät mekanismien komponentit sinänsä ole vain entiteettejä. Toisin sanoen tämä

⁸ Toki varteenotettavaa on huomioida se, että tämän monimuotoisuuden sallimisella on omat mahdolliset puutteensa, sillä määritelmästä saattaa syntyä liiankin salliva tai jopa ”tyhjä”. Esimerkiksi Phyllis Illari ja Jon Williamson (2012) eivät ole kovinkaan eksplisiittisiä sen suhteen, mitä he tarkoittavat mekanismin ilmiöstä vastaamisella (*responsibility for the phenomenon*). Termin käyttö näyttäisi siis tarvitsevan vielä tarkempaa filosofista määrittelyä tulevaisuudessa.

“tekeminen” tarkoittaa sitä, että mekanismin komponentit ymmärretään perusdualismissa aina kausaalisiksi tekijöiksi eli toiminnallisiksi entiteeteiksi (entiteetti + aktiviteetti). Voitaisiin kuitenkin sanoa, että entiteettien ominaisuudet määrittävät sen, millaisten tapahtumien tai muutoksien kohde komponentti voi olla. Aktiviteetit siis edellyttävät, että entiteeteillä on tietynkaltaisia ominaisuuksia, kuten niille ominainen koko, muoto ja rakenne; esimerkiksi vetysitoutuminen on aktiviteetti, joka on mahdollinen vain molekyylien kohdalla, jotka ovat asymmetrisesti polarisoituneita (Darden 2006, 274).

Huomion arvoista on myös se, että perusdualismin mukaan tieteen kehittyessä löydetään jatkuvasti uusia aktiviteettityyppejä. Rene Descartesin ja Robert Boylen mekanistinen filosofia tunsi esimerkiksi oikeastaan vain klassisia fysikaalisia aktiviteettityyppejä (työntäminen tai vetäminen), jotka ovat eräänlaisia produktiivisia kontaktiaktiviteetteja (Bechtel 2008, 11). Nykytiede tunnistaa aktiviteetteja, jotka eivät rajoitu vain mainittuihin klassisiin aktiviteetteihin. Näitä ovat esimerkiksi *sähkökemialliset* (kemiallinen sitoutuminen, hylkiminen), *termodynaamiset* (diffuusio) ja *elektromagneettiset* (johtaminen) (Machamer et al. 2000, 14). Näiden lisäksi on mekanistisessa kirjallisuudessa myöhemmin vielä esitetty negatiivisia aktiviteetteja (inhibitio ja estäminen) (Craver 2007, 6).

Aktiviteetti-entiteetti dualismi ei ole kuitenkaan jäänyt vaille kritiikkiä. On väitetty, että monet *perusdualistit* ovat käyttäneet entiteettitermiä ilman tarkempia rajauksia (Kaiser & Krickel 2017, 754). Tämä huomio on varteenotettava, koska useat mekanistisen teorian soveltavat entiteettitermiä ikään kuin se viittaisi vain ja ainoastaan materiaalsiin objekteihin. Oleellista on kuitenkin huomata, että entiteetti-termillä voidaan laajalaisesti viitata esimerkiksi ominaisuuksiin, prosesseihin, relaatioihin yms. seikkoihin. Näin muodoin voidaan ehdottaa, että entiteetti tulisi tulkita ontologisena kategoriana mekanistisen keskustelun kontekstissa niin, että sillä viitataan objekteihin. Vaikka tämänkään ei ole kovin selkeä raja, on se kuitenkin suuntaa antava.

Lisäksi aktiviteettien suhteen on esiintynyt samanlaisia puutteellisuksia (Kaiser & Krickel 2017, 754). Ongelmaksi muodostuu varsinkin vaatimus siitä, että mekanismien osien täytyy aina tehdä tai muuttaa jotakin aktiivisesti niin, että tämä aktiivisuus tuottaa jotakin. Mietitään esimerkiksi homeostaattisia mekanismeja, joiden tarkoituksena on *estää* tai *hillitä* muutosta (Illari & Williamson 2012, 130). Tässä konkreettisesta

esimerkistä käy se, että insuliini- ja glukagoni-hormonien tarkoitus on ylläpitää verensokeritasapainoa, eli hillitä veriplasman glukoosikonsentraation vaihtelua. Lisäksi pysähtynyt kello ei välttämättä ole mekanismi ajan kertomisen kannalta, mutta se voi olla mekanismi jonkin muun asian suhteen – esimerkiksi kilpajuoksun tuloksen kertomisen kannalta (Illari & Williamson 2012, 130). Ongelmaksi muodostuvat varsinkin passiiviset tapahtumat, jotka ovat monille biologisille mekanismeilla oleellisia (Kaiser & Krickel 2017, 754). Esimerkiksi välittäjäaineiden transmissiossa välittäjäaineet diffuusioituvat synaptisen raon halki postsynaptiseen hermosoluun, jossa diffuusioituminen on varsin ilmeinen passiivinen tapahtuma.

Aktiviteetteja on näinmuodoin ehdotettu korvattavaksi esimerkiksi tapahtuma- tai ilmenemiskäsitteellä (*occurent*) (Kaiser & Krickel 2017). Tapahtuma on siitä hyödyllinen käsitteellinen valinta, että se sallii sellaisten mekanismien ja komponenttien käsitteellistämisen, jotka eivät varsinaisesti ole aktiivisia tai tuottavia. Tapahtumat tai ilmenemiset voidaan ymmärtää paremmin jos ne asetetaan vastakkain niin sanottujen kontinuanttien kanssa (*continuants*) (Kaiser & Krickel 2017, 755). Ero näiden välillä on se, miten ne ovat olemassa ajan suhteen. Viimeiseksi mainitut ovat, hyvin karkeasti sanoen, asioita, joiden olemassa oloa määrittää jatkuvuus ajassa, kun taas ensiksi mainittujen olemassaoloa määrittää tietty aikaintervalli. Todeen perusteella seuraavat metafysiset rajaukset on syytä pitää mielessä; entiteeteillä viitataan vain objekteihin ja aktiviteetit tullaan määrittelemään tapahtumiksi. Näin ollen voitaisiin sanoa, että

Komponetti-def₂: Jokainen mekanismin komponentti on *tapahtuman alainen entiteetti*.

Mikäli mainitut huomiot pidetään mielessä, voidaan esittää seuraava esimerkki kausaalisesta mekanismista (vrt. Craver 2007, 22-25). Toimintapotentiaalisen impulsin saapuminen aksoniterminaaliin johtaa jännitteen *depolarisoitumiseen* aksoniterminaalissa, jonka vaikutuksesta jänniteherkät Ca^{2+} -ionikanavat *avautuvat*. Avautumisen johdosta Ca^{2+} -ionit *diffuusioituvat* ekstrasellulaaritalasta pre-synapsisen aksoniterminaalin intrasellulaaritalaan. Ca^{2+} -ionit *tuottavat* aksoniterminaalissa kemiallisten reaktioiden ryppään (johon tässä ei mennä tarkemmin), joiden vaikutuksesta välittäjäaineita sisältävän astian ja solukalvon välille *muodostuu* kanava, joka *sallii* välittäjäaineiden *diffuusioitua* ekstrasellulaaritalaan. Tämän kausaalisesta mekanismin kuvauksen perusteella on olemassa jokin entiteettien joukko E {pre-

synaptinen hermosolu, aksoniterminaali, Ca^{2+} -ioni, jänniteherkät Ca^{2+} -ionikanavat, ... } ja jokin tapahtumien (tässä: token) joukko T {depolarisoitua, avautua, diffuusioitua, ... }, joiden identifioiminen on oleellisessa osassa selityksen antamista. Mekanismin kuvaus ei ole vielä tällä selvä. Joukkojen E ja T sisällön identifioiminen sanoo itsessään hyvin vähän siitä, mikä on mekanismin muoto. Ilmeistä on se, että esimerkiksi edellä kuvattu transmissiomekanismi on toiminnallinen vain silloin, kun sen komponentit ovat tietyssä järjestyksessä. Toisin sanoen mekanistinen selitys ei ole yhtäläinen tapahtuman alaisten entiteettien identifioimisen kanssa. Selitys vaatii vielä sitä, että hahmotellaan se tapahtuman alaisten entiteettien konfiguraatio tai orientaatio, joka vastaa ilmiöstä.

3.2. Järjestäytyneisyys: avaruudellisuus, ajallisuus ja hierarkkisuus

Ilmiö on sitä, mitä mekanismi ”tekee” eli ylläpitää, tuottaa muutoksia, kontrolloi jne. Tätä voidaan ilmentää ”vertikaalisella” (konstitutionaalisella) ja ”horisontaalisella” (kausaalisella) akselilla. Toisin sanoen mekanismi ei ole vain osiensa summa ilmaisen ankarassa merkityksessä. Mekanismin käyttäytyy sille ominaisella tavalla, koska sen komponentit ovat järjestäytyneet tietyllä tavalla mainittujen akselien suhteen. Ilmeistä on siis se, että jokaisella komponentilla on oma roolinsa mekanismin toiminnan kannalta. Tämä rooli on kuitenkin aina suhteutettava muiden mekanismin komponenttien roolien kanssa. Yleisesti hyväksytty näyttäisi olevan se näkemys, jonka mukaan mekanismin muotoa rajoittavat sellaiset konfiguraatiosuhteet kuin avaruudelliset, ajalliset ja hierarkkiset suhteet (Machamer 2000, 3; Craver & Darden 2001; Bechtel 2008, 17; Glennan 2017, 22-25).

Avaruudelliset suhteet sisältävät sellaisia seikkoja kuten objektien paikan, muodon, koon jne. (Craver & Darden 2001, 114). Esimerkiksi välittäjäaineiden transmissiomekanismi edellyttää, että synapsien tulee olla tietyn muotoisia, kokoisia ja tarpeeksi lähellä toisiaan, jotta välittäjäaineiden kulkeutuminen niiden välillä olisi mahdollista. Oleellisia ovat myös toiseksi mainitut ajalliset suhteet (ibid.). Näitä ovat muun muassa tapahtumien kesto, järjestys ja nopeus. Tätä havainnollistaa esimerkiksi se ilmeinen seikka, että välittäjäaineiden on vapauduttava synapsiseen rakoon ajallisesti ennen kuin ne voivat kiinnittyä reseptoreihin. Lisäksi välittäjäaineet ovat vain tietyn hetken synapsisessa raossa ja niiden vapautuminen pre-synapsisesta aksoniterminaalista tapahtuu tietyissä jaksoissa. Käytännössä edeltävät kaksi suhdetta ilmaisevat sen, että jokainen mekanismin komponentti on ilmiön aika-avaruudellinen osa. Kolmas rajoittava

seikka on mekanismien hierarkkinen järjestäytyminen (ibid., 117-9). Kaikki mekanismin komponentit eivät ole välttämättä samalla tasolla. Tämä tarkoittaa sitä, että mekanismit ovat usein monitasoisia niin, että niiden joitakin komponentteja konstituoivat toiset, alemman tason (ala)mekanismit ja niiden komponentit. Esimerkiksi mainittu välittäjäaineiden transmissiomekanismi voi olla ylemmän tason muistimekanismin ”alamekanismi”.

Osiassa 2.3.2. tuotiin ilmi se, että konstitutionaalinen selittäminen koskee sitä, mikä mahdollistaa mekanismin kausaalisen kapasiteetin. Edellä määritellyt järjestyskategoriat mahdollistavat nyt tämän tarkemman analyysin. Järjestäytyneisyys on jotakin, minkä perusteella mekanistiset kokonaisuudet eroavat pelkistä aggregaattikokonaisuuksista (Craver 2007, 216; Wimsatt 1997). Aggregaatti on kokonaisuus, jonka ominaisuudet ovat ankarassa mielessä sen osien ominaisuuksien determinoimia. Muodollisemmin määritelmä voidaan antaa seuraavasti (ks. alkuperäinen muotoilu Wimsatt 1997, 376):

Aggregaatti-def:

- A1) Aggregaattikokonaisuus on ominaisuuksiltaan invariantti sen osien uudelleen järjestäytymisen suhteen, ja
- A2) ominaisuuksiltaan laadullisesti invariantti sen osien poistamisen tai uusien osien lisäämisen suhteen, ja
- A3) ominaisuuksiltaan invariantti osien hajottamisen tai purkamisen ja niiden uudelleen kokoamisen suhteen, ja
- A4) sen osien välillä ei ole vuorovaikutussuhteita (tuottavia tai inhibitorisia).

Kokonaisuuden voidaan sanoa olevan aggregaatti, jos ja vain jos se tyydyttää kaikki edeltävät ehdot. Huomion arvoista on kuitenkin se, että mainitut ehdot ovat varsin ankarat, sillä vain hyvin harva kokonaisuus tyydyttää kaikki annetut ehdot (Wimsatt 1997, 376). Esimerkiksi hiekkakasa on tyypillinen aggregaattikokonaisuus. Hiekkakasa on ominaisuuksiltaan sama hiekkakasa riippumatta siitä, miten hiekanjyvät ovat järjestäytyneet (A1). Lisäksi muutamien (saman tyyppin) hiekanjyvien poistaminen kasasta tai muutamien lisääminen kasaan ei muuta hiekkakasan ominaisuuksia kokonaisuutena (A2). Lisäksi hiekkakasa voidaan ”purkaa” pienempiin kasoihin ja myöhemmin koota jälleen uudeksi, samaksi isoksi kokonaisuudeksi ilman, että tämä uudelleen koottu kasa eroaisi ominaisuuksiltaan alkuperäisestä kasasta (A3). Vielä

viimeiseksi voidaan todeta se, ettei hiekanjyvien välillä ole varsinaisia vuorovaikutussuhteita, jotka vastaisivat hiekkakasan ominaisuuksista (A4).

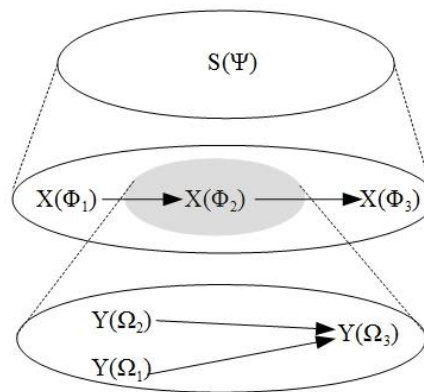
Sama ei kuitenkaan päde mekanististen kokonaisuuksien suhteen. Mekanismien voidaan sanoa olevan enemmän kuin osiensa summa sikäli kuin ne ovat heikossa mielessä (episteemisesti) emergenttejä kokonaisuuksia, joissa emergenssi on seurausta komponenttien järjestäytymisestä tietyllä; vastaavasti *vahvalla emergenssillä* tarkoitetaan sitä, että ilmiötaso ja sen totuudet eivät ole johdettavissa alemman tason seikoista edes periaatteessa (Craver 2007, 216; Chalmers 2008, §2). Ero näiden kahden emergenssin välillä on se, että ensin mainitussa ylemmän tason seikat ovat (metafyysisesti) johdettavissa alemman tason seikoista. Emergenssi seuraa kuitenkin siitä, että vaikka alemman tason seikat olisivatkin annettu, ilmenevät ylemmän tason seikat niidenkin nojalla vielä ”odottamattomina”. Odottamattomuus luo eräänlaisen episteemisen emergenssin, jossa episteeminen ”ote” ylemmästä tasosta saadaan vasta kun ollaan hahmoteltu annetun tason järjestäytyneisyys. Oleellista heikolle emergenssille on, että tämä järjestäytyneisyyden muoto voidaan johtaa kuitenkin löyhässä (metafyysisessä) mielessä niistä seikoista, jotka vallitsevat alemmilla tasoilla

Mietitään esimerkiksi muistimekanismia, joka on osittain riippuvainen hypotalamuksen hermosolullisesta rakenteesta. Muistimekanismin kannalta sillä on merkitystä, miten hypotalamuksen hermosolut ovat järjestäytyneet avaruudellisten ja ajallisten ominaisuuksiensa suhteen. Jos kaikki hypotalamuksen hermosolut järjestettäisiin uuteen, satunnaiseen järjestykseen, pelkästään tällä voidaan olettaa olevan vakavia vaikutuksia muistimekanismin toiminnallisuuteen (A1 negaatio). Lisäksi hypotalamuksen poistaminen olisi muistimekanismin kannalta käänteentekevää (A2 negaatio). On myös hyvin kyseenalaista voisiko hypotalamuksen kirurgisesti irrottaa ja uudelleen ”asentaa” väliaivoihin ilman, että tällä olisi käänteentekeviä muutoksia muistimekanismiin (A3 todennäköinen negaatio). Vielä viimeiseksi on täysin ilmeistä, että muistimekanismin osat vuorovaikuttavat keskenään (A4 negaatio). Sanottu perusteella voidaan todeta, että varsin moni kokonaisuus ja ainakin jokainen mekanistinen kokonaisuus, on vähintäänkin heikossa mielessä emergentti, sillä jokainen emergenttikokonaisuus on ei-aggregaattikokonaisuus.

Synteesinä edellä todetuista huomioista ehdotan seuraavaa mekanismien määritelmä:

Mekanismi-def: Mekanismi on tapahtuman alaisten entiteettien järjestäytynyt, ei-aggregatiivinen kokonaisuus, joka sellaisenaan vastaa ilmiöstä.

Tämä voidaan esittää myös muodollisemmin. Olkoon S ilmiö ja olkoon Ψ se aktiviteetti tai tapahtuma, jota ilmiö manifestoi. Tällöin ilmiötä voidaan nimittää notaatiolla $S(\Psi)$, joka tarkoittaa sitä, että S tekee Ψ tai S on tapahtuman Ψ alainen. Mekanismin muoto voidaan puolestaan ilmaista entiteettien joukon $E: \{X_1, X_2 \dots X_n\}$ ja aktiviteettien joukon $A: \{\Phi_1, \Phi_2 \dots \Phi_n\}$ konfiguraationa (Craver 2007, 6-7). Näin ollen on mahdollista muodostaa komponentti: $X(\Phi_1)$, joka on tapahtuman alainen entiteetti (so. X -entiteetti tekee Φ tai X -entiteetti on Φ alainen) (ibid.). Tämä kaikki voidaan esittää graafisesti seuraavasti (vrt. Craver 2007, 7):



Kaavio 1

Formaalissa mallissa kreikkalaiset aakkoset viittaavat tapahtumiin ja roomalaiset aakkoset puolestaan entiteetteihin. Nuolet ilmentävät kausaalisia suhteita, kun taas katkoviivat ilmentävät konstitutionaalista suhdetta. Lisäksi kaaviosta käy ilmi, että alaympyröiden sisältämien entiteetti- ja aktiviteettiryhmien orientaatiot konstituivat ylemmissä ympyröissä esitettyjen entiteetti- ja aktiviteettiryhmiä ja viimekädessä itse tarkasteltua ilmiötä S . Alemman tason komponentit ovat siis järjestäytyneet sisäkkäisiin hierarkioihin. Tätä voidaan tässä ilmentää seuraavalla yksinkertaisella esimerkillä.

Tarkastellaan ihmisen ruonsulatusmekanismia pintapuolisesti, jolloin S viittaa ihmiseen systeeminä ja Ψ ruoansulatustapahtumaan. Tässä tapauksessa $S(\Psi)$ koostuu monista sellaisesta osista $[X(\Phi)]$, jotka yhdessä ja tietyllä tavalla järjestäytyneenä vastaavat ruoansulatusilmiöstä. Tämän kaltaisia osia voisivat olla esimerkiksi ruuan pureskelu $[X(\Phi_p)]$, joka yhdessä sylkirauhasten erittämän syljen kanssa $[X(\Phi_s)]$ johtaa ruuan muotoutumiseen niin, että se voidaan niellä $[X(\Phi_n)]$. Useiden nielemisvaiheiden kautta ruokamassa siirtyy vatsalaukkuun. Vatsalaukussa nielty ruoka pilkkoontuu $[X(\Phi_v)]$.

Tässä vaiheessa on tarkasteltu yhden tason vaiheita horisontaalisella akselilla, mutta on myös mahdollista “zoomata” toiselle vertikaalisen akselin tasolle, jossa voidaan tarkastella $[X(\Phi_V)]$ konstituutiota. Esimerkiksi vatsalaukun pepsinogeenin aktivoituu $Y[(\Omega_A)]$ vatsalaukun mahahapon vaikutuksesta aktiiviseksi pepsiinientsyymi. Pepsiinientsyymi puolestaan konstituoii “ruoan pilkkomista” $[Y(\Omega_P)]$ sikäli kuin se osallistuu ruokamassan peptidisidoksia (proteiineja) pilkkovaan hydrolyysiin, vatsalaukun pH-arvon ollessa tarpeeksi alhainen.

Tässä ei ole kuitenkaan tarkoitus mennä tämän tarkemmin ruoansulatus mekanismiin. Oleellista on kuitenkin huomata se, kuinka monitasoisesta prosessista on kyse. Jo biotieteiden alalla konstitutionaalisia tasoja on löydettävissä makrotasolta aina molekyylitasolle asti. Eri biotieteenalat ovat usein keskittyneet eri tasojen tutkimukseen ja niillä kohdattavien ilmiöiden selittämiseen; molekyylibiologit tutkivat ja selittävät molekyylitason ilmiöitä, elektrofysiologit keskittyvät solutason sähkökemiallisiin ilmiöihin, kun taas etiologit keskittyvät eläinten käyttäytymiseen (Craver & Darden 2013, 47). Tosin usein nämä tieteenalat poikkileikkaavat. Esimerkiksi edellä mainitun ruoansulatusmekanismi on ilmiö, joka voi olla useiden alojen mielenkiinnon kohteena. Loppujen lopuksi se, mikä taso on kulloinkin selityksellisesti tärkein, on näin ollen lähinnä pragmaattinen kysymys (ibid., 48-9). Orgaanisista molekyyleistä kiinnostunut kemisti voi olla kiinnostunut pepsiinin proteiinirakenteesta ilman, että tämä selityksellinen kiinnostus ulottuu ylemmän tason ruoansulatuksellisiin mekanismeihin.

Onttisesta aspektista käsin tarkasteltuna, tähän mennessä on määritelty käsitteellisesti se, mikä on se, mitä selityksellä halutaan osoittaa (so. mekanismit). Sanalla sanoen, jos jokainen mekanismi koostuu jostakin tietyn järjestyksen komponenttijoukosta: $\{X(\Phi_i) \dots X(\Phi_n)\} \in V$ ja niiden välisistä, määrätyn kaltaisista kausaalisista suhteista, silloin ontin osoittamisen kohteena on se, että V :n sisältö ja järjestys vastaavat jotakin luonnossa olevaa sisältöä ja järjestystä. Tosin kuten edellä on todettu pelkkä puhtaasti ontin aspekti ei riitä. Tarvitaan lisäksi tapa käsitteellistää se, miten selitys eli mekanismit *valjastetaan selitykselliseksi*. Kyse on siitä, että mekanismit selittävät ilmiön aina jonkin representaation välityksellä eli *mekanistisella mallilla*. Tarkastellaan tätä tarkemmin.

4. Mallit – Se millä selitetään

4.1. Mitä mallit ovat, lyhyt yleinen kuvaus

Malli voidaan ymmärtää kolmen erillisen aspektin kautta. Tarkemmin sanoen ne toimivat *selityksellisten tekstien*, *maailman*, ja *selittäjän* välisinä välikappaleina. Tässä tullaan omaksumaan se pragmaattinen malli näkemys, jonka mukaan selittäjän selitykselliset päämäärät ovat keskeisessä osassa mallintamista (Giere 1999; 2004; Mäki 2005; 2009). Muodollisemmin mutta varsin sallivasti ilmaistuna tämä voidaan ilmaista Uskali Mäen tavoin niin, että:

Malli-def_I: “Agent A uses object M as a representative of some target system R for purpose P, addressing audience E, prompting genuine issues of resemblance to arise and applies commentary C to identify and align these components.” (Mäki 2009, 32.)

Tässä ei ole tarkoitus antaa formaalimpaa määritelmää malleista, sillä annetun määritelmän nojalla voidaan erottaa mallien keskeiset piirteet. Käytännössä A voi olla yksilöllinen tutkija, tutkimusryhmä tai jopa tiedeyhteisö kokonaisuutena (vrt. Mäki 2009, 33). Muuttujalla R puolestaan viitataan mallin kohteeseen (maailmaan), jota mallin M on tarkoitus muistuttaa. Mallin voidaankin näin ollen sanoa olevan eräänlainen korvikemaailma (*surrogate*) (Mäki 2009, 34). Korvikemaailman on tarkoitus toimia välikappaleena: mallintaja etsii informaatiota jostakin maailman kappaleen ominaisuuksista *epäsuorasti* tarkastelemalla korvikemaailman ominaisuuksia. Korvikemaailmana malli on aina itsessään jokin maailman osa niin, että tätä osaa (mallia) käytetään jonkin toisen maailman osan (mallintajan) toimesta jonkin kolmannen maailman osan (mallinnettava) representoimiseksi (Mäki 2005, 307). Näin ollen malli voi olla miltei mitä tahansa, kunhan se tyydyttää *Malli-def_I*-ehdot. Mallin ei siis tarvitse olla aina esimerkiksi yhtälö tai kaavio. Biologiassa esimerkiksi jokin organismi voi toimia jonkin biologisen mekanismin mallina (ns. malliorganismit). Näin on laita esimerkiksi *E. coli* -bakteerin suhteen, joka on varsin yleinen malliorganismi, jolla voidaan helposti ilmentää jonkin proteiinin synteesimekanismia.

Oleellista on se, että M:n ei tarvitse muistuttaa R:ää sen kaikissa suhteissa, vaan tätä muistuttavuuden astetta voidaan määrittää sen mukaan, mitä tarkoitusta (P) varten R:ää halutaan mallintaa tai millaiselle yleisölle (E) malli esitetään. Päämäärät voivat olla episteemisiä, kuten vastauksen hakeminen johonkin partikulaariseen miksi-kysymykseen, tai käytännöllisiä niin kuin silloin, kun mallista halutaan esimerkiksi

apua poliittiseen päätöksentekoon. Mallin sisällön suhteen käänteentekevää voi olla myös se, onko yleisö esimerkiksi akateeminen tai ei-akateeminen (Mäki 2009, 33.)

Mielenkiintoisen metaforan mukaan mallit ovat eräänlaisia karttoja (Giere 1999; vrt. 2004). Se, miten kartta muistuttaa kohdettaan, on riippuvainen kartan laatijan tarkoitusperistä ja siitä yleisöstä, jolle se on suunniteltu esitettäväksi. Toisin sanoen kartta on aina kohteensa *osittainen* representaatio (Giere 1999, 44-45). Esimerkiksi kaupungin topologinen metrokartta representoi kaupungin metroliikenteen tiettyjä ominaisuuksia kuten asemia ja linjoja, mutta se ei representoi ei-topologisia seikkoja, kuten esimerkiksi maaston ominaisuuksia (Giere 1999, 45-47).⁹ Näin on myös laita tieteellisten mallien suhteen. Mallit eivät siis representoi todellisuutta sen kaikessa kokonaisuudessaan vaan joidenkin ennalta asetettujen päämäärien saavuttamiseksi. Mallityypin, niin kuin karttatyyppin, valinta on näin ollen riippuvainen siitä ongelmasta, joka on tarkoitus ratkaista (Giere 2004, 750-1). Tästä voidaan johtaa myös se erittäin tärkeä huomio, *etteivät kaikki mallit ole välttämättä selityksellisiä tai kausaalisia eli mekanistisia malleja*. Tosin tulen väittämään, että on hyvät perusteet uskoa, että jokainen selityksellinen malli on mekanistinen. Näinmuodoin tässä yhteydessä tullaan keskittymään vain mekanistisiin malleihin, joiden luonnetta on syytä tarkastella seuraavaksi tarkemmin.

4.2. Mekanistiset mallit ja niiden ulottuvuudet

Jos mallit ovat jonkin asian representaatiota, niin ei liene vaikealta päätellä, että silloin mekanistiset mallit tarkoittavat mekanismien representaatiota (Glennan 2005; 2017, 66). Tarkemmin sanoen ne ovat maailman ontoksen tai kausaalisen rakenteen representaatiota. Se, mitä sanotusta voidaan päätellä, on se, että ei riitä, että malli vain muistuttaa tai ennustaa mekanismin *toimintaa*. Mekanistisen mallin täytyy muistuttaa itse mekanismia. Stuart Glennan on ehdottanut (2005, 446) seuraavaa mekanististen mallien määritelmää, jossa edellä todettu otetaan hyvin huomioon:

Malli-def₂: Mekanistinen malli koostuu i) mekanismin fenomenaalisesta kuvauksesta ja ii) mekanistisesta kuvauksesta.

Määritelmän (i)-ehto tarkoittaa sitä, että mekanistisen mallin tulee olla sellainen, että se kertoo sen, mitä mekanismi tekee; vastaavasti mekanistinen kuvaus (ii-ehto) ilmaisee

⁹ Tämä analogia on tarjottu tässä puhtaasti havainnollistavassa mielessä; sen ei ole tarkoitus esittää metrokarttoja selityksellisinä malleina.

sen, että mallin tulee tuottaa ymmärrys siitä, *miten* mekanismi tekee sen mitä se ikinä sitten sattuu tekemään (Glennan 2005, 447). Tosin on syytä tuoda ilmi, kuinka mekanistista mallia voidaan luonnehtia useista ulottuvuuksista käsin. Craver ja Lindley Darden erottelevat (2013, 59-69) neljä mekanistisen mallin ulottuvuutta: valmiustason (*completeness*); yksityiskohtaisuuden (*detail*); evidentaalisen tuen tason (*support*) ja laajuuden (*scope*). Tarkastellaan näitä ulottuvuuksia seikkaperäisemmin.

Ensinnäkin mallia voidaan analysoida sen mukaan, kuinka ”valmis” malli on. *Valmiustason ulottuvuuden* keskiössä on toisin sanoen se, missä määrin malli on relevanttien tapahtuman alaisten entiteettien konfiguraation kuvaus. Erottelu kehitemien ja skeemojen välillä on oleellinen ymmärtää, jotta välttyään siltä väärinkäsitykseltä, jonka mukaan skeema olisi kehitemmallin täydellinen vastakohta eli täydellinen selityksellinen teksti. Kehitemmallissa (sketch) kuvaus on puutteellinen, kun taas *tarpeeksi* valmis malli (skeema) on sellainen, että se sisältää pragmaattisten päämäärien näkökulmasta käsin kaikki oleelliset mekanismin vaiheet, tapahtuman alaiset entiteetit ja niiden konfiguraation (Craver 2006, 360; Craver & Darden 2013 60-61; vrt. Glennan 2005, 457). Toisin sanoen skeemamalli on eräänlainen jatkumo kehitemmallin ja täydellisen, kaikki oleelliset mekanismin vaiheet sisältävän mallin välillä.

Kyse on yksinkertaisesti siitä, että kehitemmallille ominaista on se, että se lähinnä tyydyttää vain *Malli-def₂*-määritelmän ensimmäisen ehdon. Tyypillistä kehitemmälle on siis se, ettei sen mekanistinen kuvaus ole selityksellisten päämäärien valossa tyydyttävä (Machamer et al. 2000, 18; Craver 2006, 360). Kehitemmissä mekanistinen kuvaus ei kata kaikkia oleellisia mekanismin vaihteita, mikä tarkoittaa toisin ilmaistuna sitä, että niissä kuvaus sisältää niin kutsuttuja ”mustia laatikoita” (Darden & Craver 2002, 20; Craver 2006, 360). Mustalla laatikolla tarkoitetaan sitä, että jokin mekanismin vaihe on vielä huonosti tunnettu. Informaatio joko vaiheen sisältämästä entiteetistä, tapahtumasta tai molemmista on siis empiirisen aineiston valossa puutteellinen. Mustat laatikot saatetaan kuitenkin yrittää ”täydentää” täydennystermeillä (*filler terms*), joista tyypillisimpiä ovat esimerkiksi sellaiset kuin ”aktivoi”, ”inhibitioi” tai ”aiheuttaa” (Craver 2006, 360). Tämän kaltaiset termit toisin sanoen viittaavat mekanismin tapahtumiin, mutta niissä ei täsmennetä sitä, miten ne vastaavat ilmiöstä.

Täydennystermien voidaan katsoa olevan harmittomia ja jopa hyödyllisiä tutkimuksen kannalta sikäli kuin ne osoittavat, mihin tulevaisuuden tutkimuksen tulisi keskittyä.

Pahimmassa tapauksessa ne voivat kuitenkin synnyttää illuusion siitä, että mekanistinen kuvaus olisi tyydyttävällä tasoa valmis (Craver 2007, 113). Yleisenä sääntönä voitaisiin pitää sitä, että mikäli tuleva tutkimus ei löydä täydennystermille korvaajaa, tulisi silloin malli asettaa kyseenalaiseksi (ks. Darden 2006, 286).

Valmiuden ulottuvuus on loppujen lopuksi hyvin mitään sanomaton sen suhteen, *mistä* mallissa on kyse. Yksi tapa vasta tähän kysymykseen löytyy kartoittamalla mallin *laajuuden ulottuvuutta* eli sen soveltuvuusalaa. Tämä ulottuvuus on varsinkin biotieteellisesti merkittävä, sillä harvat biologiset mallit pätevät kaikkien elämänmuotojen suhteen. Harvinaisena poikkeuksena voidaan pitää esimerkiksi James Watsonin hyvin abstraktia mallia proteiinisynteesistä ($\text{DNA} \rightarrow \text{RNA} \rightarrow \text{proteiini}$), joka representoi mekanismia, joka on (vähintäänkin periaatteessa) löydettävissä jokaisesta elävästä organismista (Craver & Darden 2013, 66). Tosin tämäkään mekanismi ei ole sovellu täydellisesti kaikkiin tapauksiin sikäli kuin RNA-viruksen infektoimat solut muodostavat omanlaisensa poikkeuksen (so. $\text{RNA} \rightarrow \text{DNA} \rightarrow \text{RNA} \rightarrow \text{proteiini}$). Toisaalta on syytä huomioda, että biologiset mallit koskevat usein mekanismeja, joiden soveltuvuusala on hyvin rajallinen, kuten esimerkiksi Huntingtonin taudin aiheuttava geneettinen virhemekanismi ihmisellä. Toisin sanoen mainitussa tapauksessa mallin soveltuvuusala koskee vain yhtä eliölajia eli ihmistä.

On syytä kuitenkin huomata, että valmiuden ja laajuuden ulottuvuudet eivät kerro sitä, missä määrin on syytä uskoa, että malli tosiasialla puhuu kohdemekanismista. Tämän takia *evidentaalinen ulottuvuus* on keskeinen. Se ilmentää sitä, missä määrin malli tukee ajatusta todellisesta mekanismista (Craver & Darden 2013, 64). Mainitun ulottuvuuden kannalta keskeistä on erottaa toisistaan kuinka-mahdollisesti-mallit (*how-possibly*) kuinka-tosiasialla-malleista (*how-actually*). Ensin mainitut ovat hypoteesin omaisia arvauksia mekanismin toiminallisuudesta, jotka ilmentävät sitä, kuinka mekanismi *saattaisi* toimia (Craver 2006, 361). Yhtäältä tämän voidaan tulkita tarkoittavan sitä, että mainitun arvauksen kohteena on luonnosta tosiasialla löytyvä mekanismi; jos arvaus on oikeassa, silloin siitä tulee kuinka-tosiasialla-malli. Tulkitsisin tämän olevan Craverin näkemys. Keskeistä Craverin näkemykselle näyttäisi olevan se, että usein näiden mallien on katsottu olevan *vain* heuristisesti arvokkaita malleja (Craver 2006, 361). Tosin tuoreessa kirjallisuudessa on esiintynyt poikkeavia tulkintoja (Koskinen 2017; Glennan 2017, 69-71). Toinen tulkintatapa on se, että kuinka-mahdollisesti-malli

saattaa olla puhtaasti mekanistisia mahdollisuuksia kartoittava. Esimerkiksi Glennan tulkitsee asiaa niin, että:

”These [how-possibly] models are not meant as conjectures about [actual] mechanisms [...] but instead are studied in their own right to understand mechanistic possibilities.” (Glennan 2017, 70.)

Näin ollen kuinka-mahdollisesti-mallit eivät ole *vain* heuristisia välineitä, vaan ne voivat olla tutkimuskohteita sinänsä. Esimerkistä käyvät muun muassa synteettisen biologian mallit, joissa voidaan kartoittaa täysin keinotekoisesti kehitetyillä kemiallisilla rakenteilla niitä vaihtoehtoisia mekanismeja, jotka olisivat voineet olla elämän perustana (Koskinen 2017, §3). Voidaan sanoa, että nämä mallit ovat puhtaasti kuinka-mahdollisesti-malleja, jotka ollaan kehitetty ja joiden kehittämistä jatketaan, vaikka tiedetään, että on olemassa niitä ”kuinka-tosiasiallisempi” DNA-malli. Craverin tulkinnasta näyttäisi seuraavan se, että nämä kuinka-mahdollisesti-mallit tulisi hylätä, koska ne voidaan korvata niitä kuinka-tosiasiallisemmilla malleilla. Tosin näyttäisi siltä, että näiden ”vaihtoehtoisen elämän aakkosien” ymmärtämisen voidaan katsoa laajentavan yleistä biotieteellistä ymmärrystä elämän luonteesta (ibid.).

Vaikka edellä todettu Craverin kritiikki on varteenotettavan, voidaan kuitenkin sanoa, että usein mekanistisen mallintamisen tarkoituksena on päästä mahdollisimman lähelle kuinka-tosiasiassa-mallia. Kuten edellä jo annettiin ymmärtää, kuinka-tosiasiassa-mallit ovat malleja, jotka ilmentävät tosiasiallista mekanismia. Usein siinä, kuinka lähelle tällaista mallia päästään, rajoitteeksi muodostuvat tutkimukselliset päämäärät (Craver & Darden 2013, 65-66). Toisin sanoen usein mallinnuksessa täytyy nojautua idealisointeihin, jotka tarkoittavat epätosien oletuksien sallimista tai ”sietämistä” mallissa. Tämä voidaan tulkita niin, että jotkin mallin muuttujien arvot tarkoituksenmukaisesti vääristetään (Kuorikoski & Ylikoski 2015; Jones 2005) tai sitä, että pyritään hiljentämään tai kontrolloimaan todellisen maailman ”melua” mallissa (Mäki 2009, 30). Biologiassa tämä saattaa tarkoittaa sitä, että esimerkiksi evoluutiota kuvaavassa mallissa jonkin ominaisuuden periytyvyysasteen (heritability) oletetaan olevan täydellinen. Yleisesti ottaen voidaan kuitenkin sanoa, että mallintajan täytyy tehdä päätös siitä, mitkä muuttujat on syytä eristää ”melusta”. Tätä päätöstä ohjaa mallintajan käsitys siitä, mitä riippuvaisuussuhteita mallilla halutaan jäljittää.

Tärkeää on, että idealisointia ei sekoiteta abstrahointiin. Viimeiseksi mainittu lukeutuu nimittäin *yksityiskohtaisuuden ulottuvuuteen*. Abstrahoinnilla viitataan tässä yhteydessä siis siihen, että malli jättää joitakin todellisuuden yksityiskohtia huomiotta.

Abstrahoinnissa kyse on kuitenkin aina asteenomaisesta yksityiskohtien huomiotta jättämisestä sikäli kuin tutkimukselliset päämäärät sanelevat sen missä määrin yksityiskohtia jätetään vaille huomiota (Craver & Darden 2013, 63). Useat näyttäisivät ajattelevan, että joskus vähemmän yksityiskohtia tarkoittaa enemmän selitysvoimaa tai selityksellistä hyveellisyyttä (Levy & Bechtel 2013; Ylikoski & Kuorikoski 2010; vrt. Raerinne 2018, §4). Mallin ei ole useinkaan tarkoitus (tai mahdollista) representoida mekanismeja sen kaikessa yksityiskohtaisessa rikkaudessaan. Siksi juuri tarvitaan abstrahointia. Sen avulla poimitaan oleelliset ilmiön taustalla olevat kausaaliset riippuvaisuussuhteet (Kuorikoski & Ylikoski 2015).

Keskeistä on myös ymmärtää, että toisin kuin evidentialisen ulottuvuuden, ei yksityiskohtaisuuden ulottuvuuden kohteena ole se, missä määrin malli tukee ajatusta tosiasiallisesta mekanismista. Sen kohteena on enemmänkin se, kuinka ”raskas” malli on kognitiivisesti. Mitä useampia yksityiskohtia mallissa on, sitä raskaampi siitä tulee. Tässä onkin oleellista huomata, että ulottuvuuksien ja selityksellisten hyveiden välillä saattaa vallita niin sanottuja kompromissisuhteita (*trade-offs*) (Ylikoski & Kuorikoski 2010, 208). Esimerkiksi merkille pantavaa on se, että mainitun kaltainen kompromissisuhde ilmenee juuri mallin yksityiskohtaisuuden ja taloudellisuuden välillä. Abstrahoinnin taso usein lisää sitä, kuinka helppoa kognitiivisilta kyvyiltään rajallisina olentoina ihmisten on omaksua ja sovittaa malli aikaisempaan tutkimukseen (Ylikoski & Kuorikoski 2010, 214–217; Bechtel & Levy 2013, 257; Raerinne 2018, 411–2). Kompleksisuus on siis usein pois mallin hedelmällisyydestä ja analyyttisestä vaikuttavuudesta. Tässä voidaan siis jatkaa edellä esitettyä kartta analogiaa toteamalla, että hyvä kartta on kätevä ja mahdollisimman helposti ja johdonmukaisella tavalla tulkittava; se ei sisällä sellaista informaatiota maastosta, joka ei ole suunnistajalle hyödyksi.

Lisäksi vaikka laajuus- ja yksityiskohtaisuusulottuvuudet ovat toisistaan riippumattomia, on niidenkin välillä eräänlainen kompromissisuhde (Craver & Darden 2013, 66–67; vrt. Ylikoski & Kuorikoski 2010, 217). Tämä käy ilmi siinä, että jos abstrahoinnin tasoa nostetaan, silloin usein soveltuvuuslaajuus kasvaa. Abstrahointi ei kuitenkaan *välttämättä* aina nosta mallin soveltuvuuslaajuutta. Sikäli kuin ne ovat

kahden eri ulottuvuuden muuttujia, on periaatteessa mahdollista, että jonkin hyvin abstraktin mallin soveltuvuuslaajuus on hyvinkin mitätön (Darden 2006, 297).

Sikäli kuin mekanistisen mallin ulottuvuudet ovat selvillä, ehdotan seuraavaksi kahta suuntaa-antavaa normatiivista muotoilua:

Idealisointinormi (IN): Epätosi oletus Q on syytä postuloida mallissa M jos ja vain jos Q postuloiminen M :ssä on välttämätöntä niiden relevanttien muuttujien eristämiseksi, joiden relevanttius on johdettavissa selityksellisistä päämääristä; tai jos Q postuloiminen lisää M :n selityksellistä hyveellisyyttä.

Abstrahointinormi (AN): Abstraktimpi yksityiskohtaisuuden aste A_1^0 on suotavampi valinta yksityiskohtaisemman asteen A_2^0 sijaan, jos A_1^0 :n nojalla selitykselliset päämäärät ovat saavutettavissa taloudellisemmin, ymmärrettävämmiin ja yhtenäisemällä mallilla M kuin A_2^0 :n nojalla.

En väitä näiden formalisointien olevan varsinaisia normien määritelmiä. Tarkoitukseni on ehdottaa annettuja normeja suuntaa-antavina, jotta tarkastelussa voidaan edetä. Tarkastellaan nimittäin seuraavaksi, miten ontisuus sopii yhteen normien kanssa.

4.3. Mekanistisen mallintamisen ja ontisuuden kritiikki

Tuoreessa kirjallisuudessa on esiintynyt kantoja, joiden mukaan ontisuuden omaksuva mekanistinen selittämiskanta ei ole yhteensopiva IN- ja AN-normien kanssa (Chirimuuta 2014; Love & Nathan 2015). On jopa esitetty niinkin radikaaleja kantoja, joiden mukaan mekanistinen selittäminen ei vastaa sitä, kuinka selitykset tosiasiaassa mallinnetaan muun muassa kognitiotieteissä. Viimeiseksi mainittujen kantojen mukaan kognitiotiede on joko ”päässyt yli” mekanistisesta selittämisestä (Chemero & Silberstein 2008) tai sen tutkimuksessa on välttämätöntä omaksua pluralistinen kanta, joka salli myös ei-mekanistiset selitykset (Chirimuuta 2014). Näitä kriittisiä kantoja näyttäisi motivoiva ajatus, että edellä esitetyt normit ovat osa tieteellistä käytäntöä, mutta että mekanistinen selittäminen, sikäli kuin se ymmärretään ontisena projektina, ei ole yhteensopiva mainittujen normien kanssa. Tarkoitukseni on seuraavaksi osoittaa, että tämä ei voi pitää paikkaansa: ontisuus on välttämätöntä hyväksyä, mikäli halutaan filosofisesti tyydyttävällä tavalla luoda käsitys siitä, mistä mallien selitysvoima on peräisin. Käsitys selityksellisistä mutta ei-mekanistisista malleista joutuisi toisin sanoen mystifioimaan selitysvoiman alkuperän.

Tarkastellaan kuitenkin ensiksi kritiikkiä, jonka mukaan mekanistinen selittäminen ei ole yhteensopiva IN- ja AN-ehtojen kanssa. Alan Love ja Marco Nathan ovat tiivistäneet hyvin idealisointia koskevan kritiikin toteamalla:

”The idealization of causal relations demonstrates that [mechanistic] models do not depict how the mechanism actually works. If actual difference makers are represented in such a way that they are not difference makers [...] mechanistic explanations appear to fail according to their own criteria.” (Love & Nathan 2015, 768).

Vastaavasti Mazviita Chirimuuta on väittänyt (2014, §3), että mekanistiseen selittämiseen näyttäisi sisältyvän hiljainen mutta epäkäytännöllinen ”mitä enemmän yksityiskohtia sitä parempi”-oletus. Kumpaakin edellä esitettyä kritiikkiä näyttäisi yhdistävän ajatus, jonka mukaan mekanistinen mallintaminen olisi vääjäämättä sitoutunut täydellisen mallin tai selityksellisen tekstin tavoitteluun. Tämä voidaan puolestaan tulkita lähinnä Craverin lujasti kannattaman ontittisuuden kritiikiksi: mikäli selityksen on tarkoitus osoittaa todellisuuden kausaalinen rakenne, silloin IN ja AN näyttäytyvät varsin ongelmallisina vaatimuksina. Näin ollen kriitikot tulkitsevat mekanistisen mallintamisen suunnan niin, että sen näyttäisi olevan aina abstrakteista kehitelmämalleista kohti täydellisempiä malleja; vastaavasti suunta kuinka-mahdollisesti-malleista näyttäisi olevan kohti kuinka-tosi asiassa-malleja. Kriitikoiden johtopäätös näyttäisi olevan yksinkertaisesti se, että nämä suuntaviivat eivät ole kestäviä. Ilmeisesti heidän mukaansa mallintamisessa tulisi omaksua ”vähemmän on enemmän”-tyyppinen mentaliteetti, jossa vähemmän yksityiskohtia ja totuudenmukaisuuden tavoittelua johtaisi kasvavaan selitysvoimaan.

Ontittisuuden on väitetty olevan ongelmallinen näkemys sikäli kuin on esitetty väitteitä sen puolesta, että monet ei-mekanistiset mallit biologiassa ja kognitiotieteessä ovat selitysvomaisia (Chirimuuta 2014; Chemero & Silberstein 2008). Näkemyksien mukaan monet tieteilijät pitävät hyvinkin abstrakteja ja idealisoituja malleja selitysvomaisina ja valmiina. Näille malleille ominaista on se, että ne ovat ei-kausaalisia malleja sikäli kuin ne tyydyttävät pitkälti vain Malli-def₂-määritelmän (i)-ehdon. Niiden suosio näyttäisi kuitenkin perustuvan niiden laajaan yleistettävyyteen ja ennustusvoimaan (Chirimuuta 2014 §5.1). Chirimuutan mukaan mekanistien perfektionistinen oletus on siis kestämaton, koska se ei kuvaa todenmukaisesti sitä,

kuinka tieteilijät todellisuudessa mallintavat selityksiä (Chirimuuta 2014, 134; vrt. Levy & Bechtel 2013).

Anthony Chemero ja Michael Silberstein väittävät puolestaan, että monet kognitiotieteilijät ovat ”luopuneet” mekanistisesta selittämisestä, sikäli kuin he ovat ottaneet haltuunsa dynaamisen systeemiteorian puhtaasti matemaattiset työkalut ilmiöiden mallintamiseksi (Chemero & Silberstein 2008, 11). Heidän mukaansa dynaamiset mallit ovat selityksellisiä mutta eivät mekanistisia, koska niitä ei motivoi kausaalisten riippuvaisuuksien ilmentäminen (ibid., 12). Tässä yhteydessä ei voida toivoa dynaamisen malliesimerkin esittelemistä sikäli kuin se vaatisi raskaan matemaattisen käsiteistön esittelemisen. Dynaamisten mallien ”prototyypinä” voidaan kuitenkin pitää differentiaalista yhtälörunkoa, kuten esimerkiksi seuraavaa lineaarista yhtälöä (Ward 2002, 64):

$$\frac{dx}{dt} = at$$

Dynaamiset yhtälöt kuvastavat toisin sanoen muutosta eli jonkin muuttujan (x) muutostahtia (rate of change) ajan suhteen (t). Käytännössä jokaisessa dynaamisessa mallissa aika on muuttujana joko eksplisiittisesti tai implisiittisesti (ibid.). Oleellista on kuitenkin ymmärtää, että mikäli mallintamisen tavoitteena on jäljittää muutostahdista vastaavat *matemaattiset riippuvaisuudet*, ei silloin mallintamisen pääasiallisena tavoitteena tai kiinnostuksen kohteena ole useinkaan *kausaalisten riippuvaisuussuhteiden* jäljittäminen. Paremmin sanoen ensin mainitut riippuvaisuudet eivät ole automaattisesti yhtäläisiä jälkimmäisten kanssa. Kannan mukaan, jota on myös kutsuttu dynaamiseksi tai komputitionaaliseksi sovinismiksi (Kaplan 2011, 341; ks. termin alkuperä Piccinini 2006), mallin selitysvaikutus ei näin ollen nouse kausaalisten riippuvaisuussuhteiden jäljittämisestä vaan (ilmeisesti) puhtaasti mallin sisäisten muuttujien välisistä matemaattisista riippuvaisuussuhteista.

Kriitikot väittävät yksinkertaisesti, että käytännössä tieteissä selityksellisinä malleina pidetään monia sellaisia malleja, jotka eivät varsinaisesti pyri jäljittämään maailman kausaalisia rakenteita. Kyse näyttäisi olevan siitä, että tieteilijät eivät kriitikoiden mukaan näytä olevan kiinnostuneita siitä, voidaanko mallin riippuvaisuussuhteet tulkita kausaalisesti. Sen sijaan näyttäisi siltä, että väitteen mukaan tieteilijät olisivat kiinnostuneita vain siitä, miten hyvin näiden riippuvaisuussuhteiden nojalla voidaan

ennustaa. Seuraavaksi tullaan kuitenkin osoittamaan, että tällä johtopäätöksellä on filosofisesti kestäättömiä seurauksia.

4.4. Miksi mallintamisen tulee olla ontista?

Osiassa 4.2. ollaan epäsuorasti jo todettu se, että mekanistiseen mallintaminen ei ole sitoutunut perfektionismiin: mekanistinen malli on aina valmis jokin annetun selityksellisen päämäärän suhteen. Itseasiassa mekanistit ovat varsin eksplisiittisesti todenneet, että ajatus täydellisestä selityksellisestä tekstistä ei lukeudu mekanistisen mallintamisen käytännöllisiin tavoitteisiin (Craver 2006; Kaplan & Craver 2011, 609–610). Tosin tämä ei tarkoita sitä, että ontisuus täytyy uhrata IN- ja AN-vaatimuksien tyydyttämiseksi. On nimittäin ehdotettu, että abstrahointeja ja idealisointeja sisältävä malli selittää ilmiön sikäli kuin se “kartoittaa malleista mekanismiin” (3M; model-to-mechanism-mapping) (Kaplan & Craver 2011, 602). Näin ollen dynaamiset mallit voivat olla selityksellisiä malleja, jos ne tyydyttävät 3M-vaatimuksen. Tämä vaatimus voidaan ilmaista seuraavasti (vrt. Kaplan & Craver 2011, 611):

Malli-def₃:

- i) Mallin selitykselliset muuttujat vastaavat kohdemekanismin identifioitavissa olevia tapahtuman alaisia entiteettejä ja niiden järjestäytyneisyyttä ja
- ii) Kaikki selityksellisinä esitetyt riippuvaisuussuhteet, jotka mallissa esitetään, vastaavat kohdemekanismin kausaalisia riippuvaisuussuhteita.

Mallien selitysvoima nousee toisin sanoen siitä, että niiden ilmentämät riippuvuussuhteet voidaan asettaa yksiin maailman kausaalisten riippuvuussuhteiden kanssa. Idealisoidut mallit ovat mielenkiintoisia, koska ne eivät ole *täysin epätosia*; ne jäljittävät joitakin kausaalisia suhteita. Samoin abstrahoidut mallit ovat mielenkiintoisia, koska ne säilyttävät *tarpeeksi* niitä yksityiskohtia, jotka ilmentävät maailman ontisia rakenteita. Tämä yksiin asettumisen ilmentäminen on se, mitä ontisuus edellyttää ja se on samalla se, mistä mallin selitysvoima nousee. *Malli-def₃* ei kuitenkaan edellytä perfektionismia, toisin kuin esimerkiksi Chirimuuta (2014) näyttäisi virheellisesti oletavan. David Kaplan ja Craver ovat varsin eksplisiittisesti todenneet (Kaplan 2011, Kaplan & Craver 2011), että perfektionismi ei sisälly annettuun määritelmään, sillä sen voidaan katsoa olevan täysin yhteensopiva abstrahoitujen mallien kanssa. Kaplan on selkeyttänyt tämän hyvin todetessaan:

“3M requires only that some of the variables in the model correspond to at least some identifiable component parts and causal dependencies among the components in the mechanism” (Kaplan 2011; 347).

Jos Craverin ja Kaplanin määritelmä hyväksytään, ei tulisi olla vaikeaa nähdä, miksi dynaamiset sovinistit ovat tyhjän päällä väittäessään, että he ovat osoittaneet tieteellisen selittämisen päässeen yli mekanistisesta selittämisestä. Voidaan nimittäin kysyä, mistä lähteestä ei-kausaalisten mallien selitysvoima nousee. Edellä todettiin, että se näyttäisi liittyvän jollakin tavalla mallin sisäisten muuttujien välisiin matemaattisiin riippuvuussuhteisiin ja ennustusvoimaan. Craver ja David Kaplan kuitenkin väittävät, että tämä vastaus ei ole tyydyttävä (Craver & Kaplan 2011; ks. myös Kaplan 2011). Tässä on syytä palauttaa mieleen se, mitä osiossa 2 todettiin eli se, että ennustusvoima ei ole yhtäläinen selitysvoiman kanssa. Edellä esitetty barometriesimerkki kuvaa tätä hyvin; malli, joka sallii myrskyn ennustamisen barometrin elohopean laskusta, on ennustusvoimainen, mutta kuitenkin selitysvoimaton. Samalla tavalla lipputanko esimerkki pätee puhtaasti matemaattisia riippuvaisuussuhteita jäljittäviin malleihin kuin se pätee peittävän lain malliin.

Vielä vakavamman ongelmallisuuden löytämiseksi, on syytä tehdä ero fenomenaalisten ja mekanististen mallien välillä (Craver 2006; Craver & Kaplan 608; Glennan 2017, 66). Yleisesti ottaen selitykselliset (mekanistiset) mallit, jotka kuvaavat maailman ontologista rakennetta, voidaan erottaa pelkistä fenomenaalisista malleista, jotka ovat vain ilmiön ja sen toiminnan kuvauksia, mutta jotka eivät sinänsä ole selitysvoimaisia. Periaatteessa voidaan sanoa, että fenomenaaliset mallit tyydyttävät edellä esitetyn Malli-def₁-määritelmän ensimmäisen (i) lausekkeen. Tärkeää on huomata, että viimekädessä dynaamiset sovinistit, jotka hylkäävät mekanistisen pohjaoletuksen ontittisuudesta, yrittävät palauttaa selitysvoimakkuuden näihin fenomenaalisiin malleihin. Puhtaasti fenomenaalisisissa malleissa piilee kuitenkin vakava ongelma, sillä sikäli kuin ne ovat hiljaisia maailman kausaalisen rakenteen suhteen, sallivat ne pahimmassa tapauksesta fiktiivisyyksien turhan sietämisen malleissa (Glennan 2017, 66; Craver 2006; 358). Ongelmana on se, että puhtaassa ei-mekanistisessa mallintamisessa ei ole riittäviä periaatteita, millä näitä mainittujen postulaattien käyttöä voitaisiin arvioida. Ainoa periaate lienee Occamin partaveitsi tai *Akaike*-kriteeri, joiden avulla voidaan valita malleista yksinkertaisempi tai taloudellisempi. Tosin nämä periaatteet eivät ole normeja,

jotka varsinaisesti kohdistuisivat fiktiivisyyksien karsimiseen sinänsä. Onttisuus puolestaan toimii tämänkaltaisen normin perustana.

Sillä, että mallintamiselta putoaa tämä normatiivinen pohja, on ikäviä seurauksia. Esimerkiksi Ptolemaioksen (fenomenaalisen) mallin nojalla taivaankappaleiden sijainti voidaan ennustaa, mutta tämä ennustuksellisuus seuraa virheellisestä maakeskeisyyden lähtöoletuksesta. Toisaalta Ptolemaioksen malli on lähestulkoon yhtä ennustusvoimainen kuin Keplerin malli, joka on lähestulkoon ensin mainitun mallin vastakohta sikäli kuin sen lähtöoletuksena on aurinkokeskeisyys. Occamin periaatteen mukaisesti Ptolemaioksen malli tosin voitaisiin hylätä. Toisaalta on ajatuskokeellisessa mielessä loogisesti mahdollista kuvitella tilanne, jossa kehitetään malli, joka hyväksyy maakeskeisen oletuksen, niin että se on ennustusvoimaisempi tämän maakeskeisyys oletuksen nojalla ja sisältää muuten vähemmän epätosia oletuksia kuin aurinkokeskeinen malli. Mikäli ontisuus hylätään, niin kuin dynaamiset sovinistit näyttäisivät tekevän, syntyy tilanne, jossa Keplerin malli ja tämä vaihtoehtoinen kuviteltu maakeskeinen malli olisivat vähintään selitysvoimaisesti yhtäläisiä. Johtopäätös on kuitenkin absurdi, mikäli halutaan pitää kiinni siitä ajatuksesta, että tiede selittää maailmaa todenmukaisesti. Tämän vuoksi tarvitaan ontisuutta.

Väitän, että dynaamiset sovinistit ovat langenneet eräänlaiseen illuusioon. Kyse on siitä, että ymmärrys mallin sisäisten muuttujien suhteista saattaa synnyttää illuusion siitä, että ymmärrys itse ilmiöstä olisi kasvanut (Kuorikoski & Ylikoski 2015, 3826). Tämän illuusion ymmärtämiseksi on syytä erottaa toisistaan maailman *ymmärtäminen mallilla* ja *mallin sisäinen ymmärtäminen* (ibid., §4, §5). Voidaan sanoa, että dynaamisilla malleilla ”leikkiminen” (muuttuja-arvojen manipuloiminen) synnyttää vain ja ainoastaan vahvan mallin sisäisen ymmärryksen. Sisäisten muuttujien ominaisuuksista ja niiden välisistä riippuvuussuhteista kasvanut ymmärrys ei automaattisesti tarkoita kuitenkaan sitä, että mallintajan ymmärrys olisi kasvanut selityskohteen todellisista riippuvaisuussuhteista (sillä nämähän ovat kausaalisia). Nämä jälkimmäiset riippuvaisuussuhteet ovat kuitenkin se, mitä viimekädessä mallilla halutaan ymmärtää tai selittää niin, että mallin ilmentämien riippuvaisuussuhteiden tulee riittävässä määrin vastata ontisia riippuvaisuussuhteita (Ylikoski 2015, 3828). Tämä on se, mitä mainittu Malli-def₃-määritelmä edellyttää.

Missään tarkastelun vaiheessa ei olla tosin tehty selkoa siitä, mitä kausaalisilla riippuvaisuussuhteilla tarkoitetaan. Yksi keskeinen keskustelu mekanistisessa filosofiassa onkin ollut se, kuinka kausaalisuus tulisi ymmärtää. Edellä ollaan esitetty hyvin abstraktilla tasolla, että mekanismi koostuu komponenteista, joiden välillä on kausaalisia suhteita, joita selitykselliset mallit representoivat. Edellä nämä suhteet merkittiin vain formaalisilla nuolilla, jotka yhdistävät komponentit. Oleellinen kysymys on kuitenkin se, kuinka nämä nuolet tulisi tulkita. James Woodwardin (2001; 2002; 2003) ns. ”kausalisuuden kontrafaktuaalinen analyysi” on ollut monien uusien mekanistien mielenkiinnon kohteena (Glennan 2002; Craver & Bechtel 2007; Craver 2007). Seuraavassa osiossa tullaan esittämään Woodwardin ajatuksia soveltaen, että nuolet voidaan tulkita kontrafaktuaalisina kytköksinä mekanismin komponenttien välillä, joiden kontrafaktuaalinen riippuvaisuus voidaan osoittaa interventioilla.

5. Manipulointavuus - Se mikä osoittaa riippuvaisuuden

5.1. Kausaaliset rakenneyhtälöt

Woodwardin mukaan kausaalinen systeemi on ilmennettävissä niin, että se rakentuu aina muuttujien varaan ja niihin arvoihin, joita nämä muuttujat voivat saada, niin että niiden välinen arvomuuttuvuus- tai eron tekevä suhde säilyy mahdollisten interventioiden alaisena (Woodward 2003, 6, 14-15). Kausaalinen järjestelmä voidaan aina puolestaan ilmaista rakenneyhtälöiden avulla, niin että jokin muuttujan arvomuuttuvuus voidaan ilmaista vain ja ainoastaan kaikkien sen suorien syiden funktiona (ibid., 42). Tarkastellaan esimerkiksi seuraavaa rakenneyhtälöä:

$$Y = f(X_1, \dots, X_n)$$

Yhtälö ilmaisee sen, että muuttujat X_1, \dots, X_n ovat muuttujan Y suoria syitä sikäli kuin Y saa arvonsa ensin mainittujen arvojen funktiona. Woodwardin kannan keskiössä on kuitenkin se, ettei tämä funktionaalisuus ilmennä vain muuttujien välistä assosiaatiosuhdetta. Kyse on sen sijaan kontrafaktuaalisesta riippuvaisuussuhteesta (ibid. 43). Toisin sanoen yhtälö tuottaa informaatiota siitä, kuinka Y -muuttujan arvot muuttuisivat, jos sen johonkin suoraan syyhyn kohdistettaisiin tietyissä olosuhteissa manipulaatio tai interventio, joka muuttaisi viimeiseksi mainitun arvoa. Kausaalinen järjestelmä voidaan tällöin ilmaista modulaarisena järjestelmänä (Woodward §7.4):

$$\begin{aligned} Y &= f_Y(X_1 \dots X_n) \\ X_1 &= f_{X_1}(X_2 \dots X_n) \\ &\dots \end{aligned}$$

Yhtälöissä funktio (esimerkissä f_Y ja f_{X_1}) ei ole koskaan tuntematon vaan on se aina spesifioitu arvomuuttuvuutta ilmentävä funktio. Lisäksi oleellista on, että yhtälöt eivät ole algebrallisia, eli käytännössä *lukusuunta on oikealta vasemmalle*. Toisin sanoen kun esimerkkijärjestelmässä " $X_1 \dots X_n$ " arvot on annettu, silloin voidaan laskea Y -muuttujan arvo. Eli syyt ovat luettavissa yhtälöstä niin, että muuttuja X_1 ja kaikki $X_1:n$ suorat syyt ovat muuttujan Y syitä. Nämä rakenneyhtälöt ilmentävät maailman kausaalisia riippuvaisuussuhteita. Näin ollen väitän, että mekanistiset mallien sisältämien riippuvaisuussuhteiden tulisi reflektoida tämän kaltaisten rakenneyhtälöiden sisältämiä riippuvaisuussuhteita. Yhtälöt antavat kausaalisien tulkinnan sille, kuinka järjestelmä käyttäytyisi mahdollisen arvomuutoksien (interventioiden) suhteen (Woodward 2003, 43). Tämä viimeiseksi mainittu huomio interventiosta on aiheen kannalta oleellinen. Sen ymmärtämiseksi on välttämätöntä tehdä selväksi se, mitä edellä mainitulla kontrafaktuaalisella riippuvaisuudella tarkoitetaan.

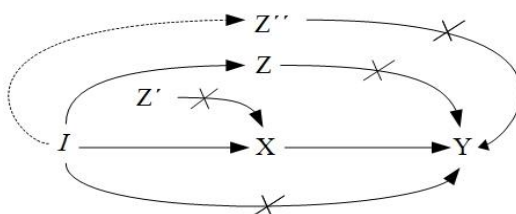
5.2. Interventionismi: kausaalisuuden kontrafaktuaalinen analyysi

Interventio voidaan määritellä woodwardilaisittain seuraavasti. Olkoon I interventio sekä X ja Y kaksi (kausaalista) muuttujaa. I on ideaalinen interventio muuttujalle X muuttujan Y suhteen, jos ja vain jos seuraavat (IV-I)-ehdot täyttyvät (Woodward & Hitchcock 2003, 13; Woodward 2003, 98):

(IV-I):

- I-i) Interventiomuuttuja I on X -muuttujan syy.
- I-ii) Kaikki kausaaliset reitit I -muuttujasta Y -muuttujaan, kulkevat X -muuttujan kautta, jolloin I -muuttuja ei vaikuta Y -muuttujan arvoihin muuten kuin X -muuttujan arvomuutoksien johdosta.
- I-iii) I -muuttuja ei korreloi minkään sellaisen muuttujan (Z') kanssa, joka on kausaalisesti relevantti Y -muuttujan suhteen niin, että se ei ole sellaisella kausaalisella polulla, joka sisältää X -muuttujan.
- I-iv) I -muuttujan vaikutuksesta X -muuttuja lakkaa olemasta kausaalisesti riippuvainen minkään toisen muuttujan (Z) arvoista.

Se, mitä tässä on yritetty todeta, on yksinkertaisesti se aiheen kannalta oleellinen seikka, että IV-I-ehdon mukaisesti määritellyt interventiot auttavat ymmärtämään, mikä tekee jostakin X:stä kausaalisesti relevantin jonkin Y suhteen. Sanottuja ehtoja voidaan ilmentää seuraavalla kaaviolla:



Kaavio 2

Kaaviossa nuolet esittävät kausaalisuhteita ja pisteillä kirjattu nuoli representoi korrelaationsuhdetta. Poikkileikatut nuolet ilmentävät puolestaan kontrolloituja kausaalisuhteita. Toisin sanoen kaavio osoittaa, että interventio muuttuja välittää kausaalisuhteensa vain yhden ja ainoan kausaalisuhteiden välityksellä ($I \rightarrow X \rightarrow Y$), ja tämä on interventionismin oleellisin ehto.

Näin muotoiltuna interventionismi on hyödyllinen käsitteellinen työkalupakki, koska sen avulla voidaan muodostaa vastauksia ”mitä-jos-maailma-olisi-ollut-toisenlainen”-kysymyksiin. Näihin kysymyksiin on puolestaan oleellista vastata, koska kausaalisuhteiden selityksen on yleensä katsottu jäljittävän kontrafaktuaalisia riippuvaisuussuhteita (Woodward 2003, 11; Craver 2007, 101; Ylikoski & Kuorikoski 2010, 205). Ilmiön selityksen on viimekädessä tuotettava informaatiota siitä, mitkä ovat ne tekijät, joista ilmiö riippuu ja osoittaa, miten ne riippuvat mainituista tekijöistä yhtäältä olosuhteiden muuttuessa että relevanttien muuttujien arvojen muuttuessa (Woodward 2003, 204). Tämä on juurin se informaatiotyyppi, mitä kontrafaktuaalisesta tarkastelusta saadaan, sillä ne ilmentävät mainittuja muutoksia. Tarkemmin sanoen selityksen vaatima kontrafaktuaalinen riippuvaisuus voidaan sanoa olevan muotoa: X selittää miksi Y, jos Y riippuu X:stä sikäli kuin, jos asianlaita ei olisi ollut X, silloin asianlaita ei olisi ollut myöskään Y (Ylikoski & Kuorikoski 2010, 204; Ylikoski 2013, 290).

Kontrafaktuaali voidaan ilmaista muuttujien välisenä eron tekijänä suhteena esimerkiksi niin, että X-muuttuja on Y:n eron tekijä, jos jossakin kontrafaktuaalisessa tilanteessa Y ottaa arvot X arvojen funktiona. Interventio on puolestaan tämän suhteen ulkopuolinen kausaalinen prosessi, joka tuottaa kontrafaktuaalin ensimmäisen lausekkeen reaktiotermien muutoksen (Woodward & Hitchcock 2003, 9). Toisin sanoen

jos Y on kausaalisen selityksen *explanandum* ja muuttujajoukko $\{X_1, \dots, X_n\}$ on *explanansin* osa, silloin se, mitä halutaan selvittää on se, päteekö rakenneyhtälön ilmaisema funktionaalisuus: $Y = f(X_1, \dots, X_n)$. Toisin sanoen funktio (f) sijoitetaan *explanansiin*, jolloin tarkoitus on havainnollistaa sitä, mitä tämä tarkoittaa Y :n ja $\{X_1, \dots, X_n\}$ joukon kannalta. Tätä voidaan puolestaan tarkentaa ilmaisemalla se kontrafaktuaalisesti: jos ideaalisen intervention johdosta $X_1 = x_1, \dots, X_n = x_n$, silloin muuttuja Y ottaa arvon näiden funktiona niin, että $Y = f(x_1, \dots, x_n)$.

Interventionismin ajatusta voidaan soveltaa hedelmällisesti edellisessä osiossa esitettyyn mallikeskusteluun. Selityksellisen mallin sisältämien muuttujien tulisi olla sellaisia, että ne voitaisiin periaatteessa tulkita sellaisiksi, että niihin voidaan kohdistaa manipulaatiota. Käytännössä tämä tarkoittaa sitä, että mallin tulisi mahdollistaa mahdollisimman monia kelvollisia *mitä-jos*-päätelmien tekemisen. Mallin selityksellisen syvyyden tai voiman mitaksi voidaan väittää juuri sitä, kuinka laajan joukon *mitä-jos*-päätelmiä se mahdollistaa (Kuorikoski & Ylikoski 2015). Tosin tämä edellyttää mallia siitä, että on olemassa jokin hyvin määritelty tapa manipuloida jotakin sen muuttujista (vrt. Woodward 2000, 202; 2003, 111-2, 247). Siispä jos ei ole mahdollista muodostaa käsitystä siitä, mitä muuttujan arvon muuttaminen tarkoittaa, so. antaa sille kausaalista tulkintaa, silloin on syytä epäillä mallin selitysvoimaa. Näinmuodoin voidaan väittää, että esimerkiksi puhtaasti pelkkiä matemaattisia riippuvaisuussuhteita jäljittävät dynaamiset yhtälöt ovat selitysvoimaltaan kyseenalaisia sikäli kuin niiden suhteen on vaikea antaa kausaalista tulkintaa sille, mitä mallien sisäisten muuttujien manipuloiminen tarkoittaa.¹⁰

Eron tekevyyden paikantamisessa on siis kyse siitä, että hahmotetaan se, milloin, missä määrin tai minkä ehtojen suhteen yhden muuttujan arvomuutokset johtavat muutoksiin jossakin toisessa muuttujassa ensin mainitun arvomuutoksen funktiona. Nämä suhteet puolestaan ymmärretään representoivan mekanismin sisäisiä suhteita. Sanotun perusteella kysymys: ”miten hermosolun depolarisoituminen selittää toimintapotentialiaalin?”, kääntyy kysymykseksi: ”miten toimintapotentialiaali riippuu hermosolun depolarisoitumisesta?” Tämä voidaan vielä asettaa kontrastiivisesti niin, että ”miksi depolarisoitunut hermosolu laukaisee toimintapotentialiaaleja, kun taas

¹⁰ Tämä kausaalisuuden tulkinnallisuus ei koske ainoastaan kognitiotieteessä esiintyviä dynaamisia malleja, vaan sama voidaan yleistää koskemaan kaikkia tieteenaloja. Vastaavan kaltaisia argumentteja on esitetty esimerkiksi ekologisten mallien kohdalla (ks. Raerinne 2011 122-138).

depolarisoitumaton ei tee näin?” Riippuvaisuus jonkin asteisesta depolarisaatiosta voidaan osoittaa kontrafaktuaalien nojalla esimerkiksi niin, että

(5.2.2): Jos tilanne olisi w ja hermosolu depolarisoitaisiin (interventio) yli kynnysarvojännitteen (n. $-50/-55\text{mV}$), silloin hermosolu laukaisisi toimintapotentiaalin.

Tilanteella w tarkoitetaan tässä tilannetta, jossa tietty joukko olosuhdemuuttujia pidetään vakiona (ympäristön pH-arvo, lämpötila jne.). Oleellista on huomata, että hermosolun jännite voidaan kokeellisessa mielessä pitää vakiona, mikä on manipulatiivis-kontrafaktuaalisen tarkastelun anti. Viimeistään tällöin voidaan osoittaa, miksi toimintapotentiaali on riippuvainen depolarisaatiosta. Sanalla sanoen:

(5.2.3): Jos tilanne olisi w ja hermosolun kalvojännite pidettäisiin vakiona alle kynnysarvojännitteen (esimerkiksi sen lepojännitteessä -70mV), silloin mikään hermosolun kohdistettu interventio ei tuottaisi toimintapotentiaalia.

Näin ollen voidaan sanoa, että tilanteen ollessa w hermosolun depolarisaatio on kausaalisesti relevantti eli eron tekevä toimintapotentiaalimekanismille.

Tulee huomata myös se, että eron tekevien suhteiden tulee päteä useissa aktiivisissa, että passiivisissa kontrafaktuaalissa tilanteissa (tarkempi ero näiden kahden välillä ks. Woodward 2000, 211, etenkin alaviite 2). Aktiiviset kontrafaktuaalit ovat käytännössä interventioiden synnyttämiä kontrafaktuaaleja (X -muuttujan arvon mukauttaminen), kun taas passiiviset kontrafaktuaalit ovat olosuhdemuutoksien synnyttämiä kontrafaktuaaleja. Nämä kaksi nostavat esille kaksi tärkeää ehtoa; sanalla sanoen se aktiivisten kontrafaktuaalin laajuus, jonka suhteen yleistys pätee, nostaa esille invarianssikriteerin ja vastaavasti passiivisten kontrafaktuaalien laajuus nostaa esille stabiiliuskriteerin. Nämä kaksi, invarianssi ja stabiilisuus, eroavat oleellisesti toisistaan (Woodward 2001, §8). Invarianssilla tarkoitetaan sitä, missä määrin eron tekevä suhde ei ole herkkä X -muuttujan intervention tuottamille arvomuutoksille. Esimerkiksi X ja Y välinen eron tekevä suhde on invariantti, jos ja vain jos se ei ”rikkoudu” joissakin hypoteettisissa olosuhteissa, joissa X arvo muuttuu jonkin intervention johdosta (Woodward & Hitchcock 2003, 15). Vastaavasti yleistystä voidaan pitää stabiilina, jos se pätee useiden olosuhdemuuttujien muutoksien suhteen. Lyhyesti on vielä todettava se, että edeltävässä osiossa esitelty soveltuvuuslaajuus (scope) ei ole sama asia kuin

stabiilius (ks. Raerinne 2011, 150-152). Viimeiseksi mainittu koskee nimittäin mallin (modaalista) pätevyyttä *mahdollisten* olosuhdemuutoksien suhteen, kun taas ensin mainittu koskee (ei-modaalista) *tosiasiallista* pätevyyttä menneisyydessä ja nykyisyydessä.

Esimerkeissä 5.2.2 ja 5.2.3 mainittu w on oleellinen määritellä, sillä sen avulla voidaan määrittää se olosuhteiden laajuus, jonka suhteen mekanismi on stabiili, sekä se invarianssiala, jonka suhteen eron tekevä suhde muuttujien välillä ei rikkoudu.

Mekanistisessa teoriassa nämä molemmat ovat oleellisia, sillä useat (esimerkiksi biologiset) mekanismit ovat ”hauraita” (vrt. Glennan 1996; Craver 2007, 68, 72, 200).

Toisin sanoen ne eivät ole yleistettävissä laajojen kontrafaktuaalisten laajuuksien suhteen sikäli kuin olosuhdemuuttajat ja sisäiset muuttajat luovat rajoituksia yleistettävyydelle.

Sanalla sanoen se, mitä halutaan sanoa, on se, että hermosolun depolarisointi johtaa toimintapotentialin laukaisemiseen useissa erilaisissa kontrafaktuaalisissa tilanteissa (niin aktiivisissa kuin passiivissa), jolloin kyse ei ole vain sattumasta.

Tätä voidaan ilmentää esimerkiksi niin, että hermosolun aksonin pituus, hermosolun sijainti tai depolarisoitumisen aste (mikäli se ylittää kynnysarvojäännitteen) ei pitäisi vaikuttaa tietyissä asteissa tai sijainneissa jne.

esimerkkien 5.2.2 ja 5.2.3 yleistettävyyteen. Johtopäätös on kuitenkin se, että invarianssi ja stabiilius eivät edellytä, että eron tekevän suhteen tulee olla järkkymätön kaikkien mahdollisten kontrafaktuaalien suhteen. Mekanistisen selittämisen kannalta käytännössä oleellista on siis määritellä ärsyke-ehdot, välilliset ehdot ja taustaehtot.

Ärsyke-ehtojen tarkoituksena on määritellä se ärsykearvojen skaala, jonka suhteen mekanismi toimii normaalisti (Craver 2007, 99; vrt. Woodward 2003, 249-250). Tämän skaalan ulkopuolisten ärsykearvojen suhteen mekanismia koskevan yleistyksen ei voida olettaa olevan invariantti. Periaatteessa ärsyke-ehtojen kohteena on manipulaatiosuhteen itsenäinen muuttuja (Esimerkiksi $X \rightarrow \dots \rightarrow Y$ suhteessa itsenäinen muuttuja on X).

Manipuloitavuus edellyttää sitä, että X -muuttujaan tulee kohdistua *riittävä* ärsyke, joka toimii mekanismin laukaisijana (vrt. Woodward 2003, 250). Oleellista on, että tämä ”riittävyys” voidaan määritellä selvästi. Tästä päästään siihen, että yleistyksen ei tarvitse päteä kaikkien ärsykearvojen suhteen, koska mekanismi ei välttämättä ole toiminnallinen kaikkien ärsykesyötearvojen suhteen. Skaalan ulkopuoliset ärsykearvot saattavat johtaa esimerkiksi siihen, että mekanismin komponentti vaurioituu, eikä se silloin ilmennä normaalia toiminallisuuttaan. Vastaavasti ärsyke voi olla niin heikko,

ettei se riitä aktivoimaan mekanisme. Esimerkiksi nykytietämyksen valossa hermosolun täytyy depolarisoitua vähintään $-50/-55\text{mV}$, jotta se voisi laukaista toimintapotentialin (Craver 2007, 209). Toisin sanoen kaikki muutokset kalvojännitteessä eivät tuota muutoksia toimintapotentialissa. Näin ollen voidaan sanoa, että ärsyke-ehtojen tarkoituksena on havainnollistaa eräänlaista mekanistista ”vipupistettä” (vrt. Craver 2007, 202), joka määrittää sen, että $Y = f(x_1, \dots, x_n)$ pätee luotettavasti vain ja ainoastaan jonkin määrätyn arvojakauman $\langle x_1, \dots, x_n \rangle$ suhteen. Tällöin mainittua arvojakaumaa voidaan pitää vipupisteenä.

Välillisten ehtojen tarkoituksena on osoittaa se välillisten muuttujien kelpoisuus, jonka suhteen mekanistinen yleistys pätee (Tämä viittaa toisin sanoen reitin $X \rightarrow \dots \rightarrow Y$ muuttujien välisiin muuttujiin) (Craver 2007, 99; Woodward 2003, 248). Välilliset ehdot eroavat siis edellä esitetyistä ärsyke-ehdoista sikäli kuin ne eivät koske itsenäistä muuttujaa (X), vaan sen sijaan kausaalisella reitillä esiintyviä riippuvaisia muuttujia. Näiden riippuvaisten muuttujien tulee olla mekanismin toiminallisen prosessin etenemisen ”kelpuuttavia”. Esimerkiksi välittäjäainetransmissiossa tämän kaltainen ehto ottaa huomioon sen, että transmissio on mahdollinen vain kun välittäjäainepitoisuus¹¹ on tarpeeksi suuri tai vastaanottavan hermosolun reseptoreita on tarpeeksi aktivoitavissa.

Lisäksi oleellista on huomata se, että mekanistinen selitys on yleistettävissä vain tiettyjen *taustaehtojen täytyessä* (Craver 2007, 99; vrt. Woodward 2003, 248). Nämä ehdot koskevat niitä muutoksia, jotka ilmenevä yleistyksen ulkopuolisten muuttujien suhteen. Biologisten mekanismien kannalta tärkeitä taustamuuttujia ovat esimerkiksi ympäristön pH-arvo, lämpötila ja vapaan energian saatavuus. Esimerkiksi vatsalaukun ruokaa hajottavat peptiinientsyymit lakkaavat toimimasta normaalisti, jos vatsalaukun pH arvo laskee tai nousee yli tietyn arvon. Kun taustaehtojen laajuus on määritelty, silloin mekanistinen yleistyksen stabiilisuuden aste on annettavissa, mikä heijastaa sen luotettavuutta. Yleistys koskien peptiinientsyymien ja proteiinien pilkkomista vatsalaukusta on siis luotettava, kun siinä ollaan esimerkiksi otettu huomioon se pH-arvo, jonka suhteen entsyymi on toiminnallinen. Tärkeää on huomata, että myös sattumanvaraisuudet tai korrelaatio-suhteet voivat pysyä stabiileina tiettyjen

¹¹ Välittäjäainepitoisuus ei ole itsenäinen muuttuja sikäli kuin se on välittäjäainetransmissiomekanismissa *riippuvainen* depolarisaatiosta. Näinmuodoin depolarisaatiota voidaan merkitä itsenäisellä X:llä ja välittäjäainepitoisuutta jollakin mainitusta riippuvaisella muuttujalla O, joka asettuu muuttujien kausaalisella reitillä muuttujien X ja Y välille.

olosuhdemuutoksien suhteen (Woodward 2001; 2003, 248). Woodward näyttäisikin ajattelevan (2001, 4-5), että tästä syystä invarianssi on etusijalla stabiilisuuden nähdessä, kun kyseessä on yleistyksen selitysvoima. Tosin on esitetty, että stabiilius on tärkeä yleistyksen ominaisuus sikäli kuin se ilmentää yleistyksen luotettavuuden astetta; mitä laajempaa passiivisten kontrafaktuaalien joukkoa yleistys tukee, sitä luotettavampi sen voidaan sanoa olevan (Raerinne 2011, 170-172).

Todetun perusteella voidaan sanoa, että interventionismi on tapa, jolla mekanismin kausaalista verkostoituneisuutta voidaan jäljittää. Sanalla sanoen, kun tiedetään mekanismin interventiosuhteet, voidaan siitä johtaa se, mikä on selityksen kannalta relevanttia. X on kausaalisesti, ja siten selityksellisesti, relevantti ilmiölle Y, jos on olemassa sellainen tilanne w, jonka suhteen Y arvot ovat riippuvaisia X arvoista (vrt. Woodward & Hitchcock 2003, 11). Tarkastellaan seuraavaksi hieman seikkaperäisemmin, minkä vuoksi interventionismin soveltaminen kausaalisten mekanismien analyysiin on hedelmällistä.

5.3. Mikä tekee interventionismista houkuttelevan?

Mitä hyötyä tästä interventionismin käsitteellisestä työkalusta on tieteellisen selittämisen teorialle? Vastauksen ymmärtämiseksi syytä on palauttaa mieleen ne edellä osoitetut huomiot (S1-S5), jotka ilmentävät selityksen kausaalista luonnetta. Lienee aiheellista tarkastella, kuinka interventionismi tyydyttää nämä viisi kohtaa.

Tarkastelun perusteella tulisi olla ilmeistä, että ei-kausaalinen korrelaatio voidaan selittää pois interventionismin avulla, koska korrelaatiot eivät ole toistensa eron tekijöitä – so. manipuloimalla yhtä korrelaattia ei voi tuottaa arvomuutoksia toisessa korrelaatissa (Woodward 2003, 253; Craver 2007, 102). Tarkastellaan tätä esitetyn barometri-esimerkin valossa kysymällä, miten epäaito syy-seuraussuhde barometrin lukeman ja myrskyn välillä voitaisiin *selittää pois*. Vastaus kysymykseen on varsin yksinkertainen; barometriä manipuloimalla ei voida tuottaa muutoksia siinä, milloin myrsky ilmenee, kun ilmanpaine pysyy vakiona. Näinmuodoin S1-ehto on yhteensopiva interventionismin kanssa.

Voidaan helposti todeta myös se, ettei interventionismi edellytä, että jonkin seikan manipuloiminen tekee ilmiöstä todennäköistä (Craver 2007, 102). Se, mikä interventionismin nojalla on oleellista, on se, että manipulaatio vaikuttaa ilmenevien seikkojen todennäköisyyteen. Jos X on selityksellisesti relevantti Y:lle, silloin tulisi olla

niin, että jos asianlaita olisi X, niin todennäköisyys sille, että asianlaita olisi Y toinen kuin mitä se olisi, jos asianlaita ei olisi X. Sanalla sanoen X:n ei tarvitse tehdä Y:stä *todennäköistä* tai odottavissa olevaa, vaan riittää, että X on eron tekevä Y:n todennäköisyyden suhteen. Siispä interventionismin soveltaminen selitysteoriaan tyydyttää kohdan S2.

Lisäksi ehto S3 tulee tyydytykseksi, koska irrelevantteihin seikkoihin kohdistetut interventiot eivät tuota muutoksia ilmiön relevanteissa seikoissa (Woodward 2003, 361). Jos siis osion edellä esitetyn vastaesimerkin 2.1.2.3. kuvaamassa tapauksessa manipuloitaisiin X:n ehkäisy pillereiden syömistä, ei tällä manipulaatiolla voitaisi katsoa olevan vaikutusta X:n raskaaksi tulemiseen.

Vielä voidaan todeta interventionismin nojalla se, että selittävä tekijä on aina se, mitä voidaan manipuloida, ja selitettävä tekijä on se, johon ensin mainitun tekijän manipulointi vaikuttaa¹². Oleellista on kuitenkin huomata, että seurausta manipuloimalla ei voida tuottaa arvomuutoksia sen syyssä. Näin ollen lipputangon varjoa manipuloimalla ei voida tuottaa muutoksia lipputangon pituudessa (Woodward 2003, 361), tai manipuloimalla hermosolua hetkenä t_2 ei voida vaikuttaa siihen tilaan, jossa hermosolu oli hetkenä t_1 (vrt. Craver 2007, 102). Interventionismin avulla voidaan siis hahmottaa selkeästi se, mikä selittää minkäkin. Todetun perusteella S4 ehto tulee tyydytetyksi.

Biologisten selityksien kannalta interventionismi tarjoaa myös hedelmällisen tavan käsitteellistää poissaolevuudet (*ommission*) (vrt. Woodward 2003, §5.11; Craver 2007, 104). Poissaolevuuksiin viitataan varsin usein biologisissa selityksissä. Tätä voidaan havainnollistaa sillä, kuinka Langerhansin saarekkeiden alfasoluista erittyy sokeriaineenvaihduntamekanismiin osallistuvaa glukagonihormoonia. Hyvin pelkistetty selitys on se, että glukagonia erittyy haimasta veriplasman glukookonsentraation ollessa alhainen (*hypoglycemia*, alhainen verensokeritaso). Toisin ilmaistuna glukooksin tietyn asteinen ”poissaolevuus” on glukagonin erittymisen syy. Tämä voidaan käsitteellistää niin, että verensokeritasoa manipuloimalla voidaan vaikuttaa siihen,

¹² On oleellista huomata, ettei manipulaation tarvitse olla inhimillisen agentin suorittama, aktuaalinen manipulaatio (Woodward 2001, 4). Kyseessä voi olla pelkästään teoriassa suoritettava manipulaatio. Periaatteessa manipuloitavuus on ehkä parempi ymmärtää interventio-termin avulla, joka ei sisällä samanlaista inhimillisyyteen viittaavaa ulottuvuutta kuin manipulaatio-termi. Näin muodoin myös ”luonnon suorittamat” interventiot, kuten alla seuraavassa osiossa esitettävässä esimerkissä esiintyvä sydänkohtaus, voidaan ymmärtää interventioksi.

miten glukagonia erittyy. Vipupiste voitaisiin määritellä esimerkiksi niin, että kun plasman glukoositasot alkavat olemaan alle 3.7 mmol/l, silloin negatiiviset arvomuutokset (so. interventiot) glukoositasoon vaikuttavat glukagonin erittymisen. Tarjotussa esimerkissä on kyse asteenomaisesta puuttumisesta, mutta interventionismin avulla voidaan käsitteellistää myös täydellinen poissaolevuus. Näin on esimerkiksi laita NMDA-reseptorien aktivoitumisen suhteen (Craver 2007, 80). NMDA-reseptorin aktivoituminen tarkoittaa sitä, että sen proteiinkanava aukeaa, ja se sallii siten Ca^{2+} -ionien diffuusioitua post-synapsiseen soluun. Tosin inaktiivisiin NMDA-reseptoreihin sitoutunut Mg^{2+} -ioni tukkii mainitun proteiinkanavan, mikä estää Ca^{2+} -ionien sisään virtauksen. Interventionismin nojalla voidaan sanoa siis, että Mg^{2+} -ionin ”poissaolevuus” on Ca^{2+} -ionivirran kannalta eron tekevä. Näin muodoin se on myös sisään virtauksen syy, joka käy selvästikin ilmi siitä, että manipuloimalla NMDA-reseptorin Mg^{2+} -ionia, voidaan vaikuttaa Ca^{2+} -ionivirtaan.

Todetun perusteella poissaolevuudet ovat mahdollisten manipulaatioiden kohteita, mikä tekee niistä myös selityksellisesti relevantteja eron tekijöitä. Mikäli kausaalinen relevanssi ymmärretään kontrafaktuaalisen analyysin mukaisesti, voidaan sitä soveltaa silloin laajemmin kuin niin kutsuttuja, kenties klassiseksi muodostuneita, kausaalisuuden ”transmissioteorioita”. Transmissioteorioissa kausaalisuutta ei voida ymmärtää komparatiivisena eron tekevänä suhteena, sillä niissä kausaalisuhde on ilmaistu jonkin aiheuttavan syyn ja sen seurauksen välisenä fysikaalisena yhteysprosessina (Woodward 2002, §4; 2003, 4, 35). Toisin muotoiltuna mainituissa kannoissa kausaalisen prosessi on joko kykyä välittää jokin fysikaalisen ”merkki” (Salmon 1984, 139-147; Salmon 1998, 15-16) tai vastaavasti kykyä välittää syyobjektin ja seurausobjektin välillä jotakin fysikaalista kvantiteettia (energiaa, massaa, varausta jne.) (Salmon 1998, 18-19; Dowe 2000, 90). Näin ollen kausaaliset vuorovaikutukset olisivat pitkälti klassisen biljardipallojen törmäyksen kaltaisia yhteysprosesseja; valosäteen eteen tuotu punainen filtteri ”merkitsee” valosäteen, jolloin tämä filtteriin ja säteen välinen kausaalinen yhteys säilyy valosäteessä filtteriin jälkeen. Vastaavasti biljardipallot vaihtavat törmätessään liike-energiaansa. Tässä yhteydessä ei ole kuitenkaan tarkoitus mennä tarkemmin näihin kantoihin. Oleellista on kuitenkin huomata se kontrasti, joka vallitsee niiden ja edellä esitetyn kontrafaktuaalisen analyysin välillä. Interventionismia soveltava kontrafaktuaalinen analyysi on kausaalisen relevanssin suhteen sallivampi, koska se ei edellytä mitään välittävää

yhteyttä syyn ja seurauksen välillä. Edellisessä kappaleessa esitettyjen poissaolevuusesimerkkien nojalla tulisi olla kuitenkin selvää, ettei poissaolevuus ole ongelma kontrafaktuaaliselle analyysille, koska syyn ja seurauksen välillä ei odoteta olevan mitään fysikaalista välittävää suhdetta (Woodward 2002, 372; Craver 2007, 78).

Tosin vielä ei ole selvää, kuinka interventionismin avulla voidaan tyydyttää konstitutionaalisen selittämisen kannalta oleellinen S5-ehdo, eli osoittaa, että ilmiön toiminnalliset osat ovat ilmiön selittäviä. Tämä onkin mekanistisen selittämisen haasteellisimpia aiheita, josta ei näytä olevan minkäänlaista konsensusta. Woodwardin (2003) interventionismia ei ole alun perinkään kehitelty konstitutionaalisen relevanssin teoriaksi, vaan pikemminkin kausaalisen relevanssin. Toisin sanoen ei ole ollut täysin selvää, onko ylemmän tason ominaisuuksilla, jotka eivät ole identtisiä niitä realisoivien ominaisuuksien kanssa, kausaalista voimakkuutta (kts. Woodward 2015). Monet mekanistit ovat kuitenkin pyrkineet soveltamaan interventionismia konstitutionaaliseen analyysiin, jonka vuoksi onkin aiheellista siirtyä näiden sovellusyrityksien seikkaperäisempään tarkasteluun.

6. Interventionismin soveltaminen konstitutionaaliseen analyysiin

6.1. *Mutuaalinen manipulaatio*

Millä perusteella mekanismin relevantit komponentit voidaan käsitteellistää? Oleellista on nimittäin huomata, että myös konstitutionaalisten selityksien on katsottu jäljittävän kontrafaktuaalisia riippuvaisuus- eli eron tekeviä suhteita (Ylikoski 2013, 278). Interventionismin nojalla intuitiivisesti helpolta vastaukselta kysymykseen, mikä tekee mekanismin komponentista relevantin mekanismin kokonaisuudelle, saattaisi siis vaikuttaa seuraava; $X(\Phi_i)$ on relevantti $S(\Psi)$ suhteen, jos ja vain jos $X(\Phi_i)$ ja $S(\Psi)$ välillä vallitsee arvomuuttuvuus eli eron tekevä suhde. Tämän eron tekevyys voitaisiin puolestaan paljastaa interventioiden avulla kontrafaktuaalisessa analyysissä. Näin ollen mekanistinen relevanssiehto olisi puhtaasti interventionismin avulla muotoiltavan kausaalisen relevanssiehdon sovellus:

Jos ideaalisen intervention vaikutuksesta komponenttimuuttuja $X(\Phi_i)$ ottaa arvon φ_1 ($X(\Phi_i) = \varphi_1$) ja ilmiötason muuttuja $S(\Psi)$ ottaa arvon sen funktiona niin, että ($S(\Psi) = f(\varphi_1)$), silloin komponenttimuuttuja $X(\Phi_i)$ on relevantti ilmiötason muuttujalle $S(\Psi)$ (vrt. Craver 2007, 155).

Yllä esitetty ehto ei ole kuitenkaan sopiva konstitutionaalisen selityksen relevanssiehdoksi (Craver 2007, 156), mikä tullaan osoittamaan seuraavaksi tarkemmin. Esitetyn ehdon voitaisiin katsoa olevan sovellettavissa vain kausaalisten selityksien suhteen. Edellä esitetty hermosolun depolarisaatiota käsittelevä esimerkki on tyypiesimerkki kausaalisesta mekanismista, jossa relevantit seikat voidaan määrittää puhtaasti kausaalisen relevanssin nojalla. Kausaalinen relevanssi ei ole kuitenkaan yhtäläinen konstitutionaalisen relevanssin kanssa, sillä ensin mainittu jäljittää kausaalisia riippuvaisuuksia, kun taas jälkimmäinen tasojen välisiä riippuvaisuuksia. Alkuperäisen (woodwardilaisen) interventionismin suora soveltaminen konstitutionaaliseen selittämiseen ei siis käy päinsä. Suora soveltaminen tarkoittaisi sitä, että relevanssi määräytyisi vain alhaalta ylös, bottom-up-interventiosuhteiden nojalla, mikä on ongelmallinen käsitys, kuten tullaan seuraavaksi esittämään.

Ensimmäiseksi huomattavaksi ongelmaksi muodostuvat *epäsuorat vaikutukset*. Aivovauriot ovat usein kiinnostavia tutkimuksen kohteita, koska ne ovat tyypillisiä bottom-up-interventioita. Vaurio jollakin aivoalueella saattaa tuottaa muutoksia jossakin tietyssä kognitiivisessa mekanismissa, josta parhaassa tapauksessa voidaan muodostaa hypoteesi, että kyseisen aivoalueen toiminta on mainitun kyvyn komponentti (Bechtel 2008, 42). Aivot ovat kuitenkin holistinen järjestelmä; vaurio yhdellä aivoalueella voi johtaa laajoihin muutoksiin toisilla aivoalueilla, jolloin pelkän paikallistuneen aivovaurion perusteella on mahdotonta paikantaa jonkin kognitiivisen kyvyn *relevanttia* komponenttia varmuudella (ibid.). Tästä voidaan siis päätellä, että interventio jonkin irrelevantin muuttujan $Y(\Phi_i)$ suhteen voi tuottaa muutoksia jossakin relevantissa $X(\Phi_i)$ -muuttujassa, joka on ilmiön $S(\Psi)$ komponentti.

Sanotun lisäksi toisenlaisen ongelman muodostaa mahdollinen järjestelmän sisäinen *kompensatio*. Joskus konstitutionaaliseen pohjaan kohdistettu interventio ei tuota havaittavia vaikutuksia mekanismissa. Kyse voi olla siitä, että järjestelmä kompensoi johonkin komponenttiin $X(\Phi_i)$ kohdistetun muutoksen (Craver 2007, 150; Bechtel 42; Craver & Darden 2013, 188. Esimerkiksi jonkin aivoalueen vaurioituminen voi johtaa siihen, että toinen aivoalue alkaa suorittaa vaurioituneen alueen normaaleja funktioita, jolloin intervention ei tuota havaittavia vaikutuksia (Bechtel 2008, 42). Toinen tyypiesimerkki tästä on toleranssi (Craver 2007, 150). Kohdemekanismin konstitutionaalisuutta voidaan esimerkiksi pyrkiä havainnollistamaan lääkeainevaikutuksien avulla, joka tuottaa muutoksia relevanteissa komponenteissa ja

siten kokonaisuudessa. Ongelmana tässä on kuitenkin se, että lääkeaineen suhteen voi kehittyä toleranssi, jolloin intervention vaikutus jää havaitsematta. Näin ollen jonkin relevantin $X(\Phi_i)$ komponentit stimuloiminen ei välttämättä johda muutoksiin ilmiötason $S(\Psi)$ muuttujassa (Craver 2007, 148, 150).

Tieteellisen käytännön valossa kompensatio ongelma ei ole välttämättä ole niin dramaattinen, kuin filosofiset esimerkit antavat ymmärtää. Voitaisiin nimittäin ehdottaa, että manipulatiivinen vaikutus, jonka nojalla relevanssi havaitaan, voidaan yrittää havaita *ennen* kompensatiovaikutusta esimerkiksi kontrolloimalla kompensatiovaikutuksen ilmenemistä. Tämä on ehkä pätevä puolustus tieteellisen käytännön kontekstissa. Väitän, että filosofiselta kannalta tarkasteltuna tämä on huono puolustus, sillä ei olemassa sellaista loogista taetta, että manipulaatio- ja kompensatiovaikutuksen välillä on välttämättä ajallinen ero. Toisin sanoen voitaisiin kehitellä sellainen (absurdi) ajatuskoe, jossa manipulaatio- ja kompensatiovaikutus alkavat samanaikaisesti; esimerkiksi niin, että ”noiduttu” lääkeannos nostaa nautittuna välittömästi subjektin lääketoleranssin sen tasoiseksi, ettei ole mahdollista havaita annoksen manipulatiivisia vaikutuksia.

Toisaalta vielä voitaisiin ehdottaa, että kompensatio ja epäsuorat vaikutukset voitaisiin poissulkea tukevapohjaisella (*robust*) evidenssillä eli evidenssillä, jota tuotetaan useista toisistaan riippumattomista koeasetelmista käsin. Käytännössä tämä tarkoittaisi muun muassa stimulaatio- ja häirintäinterventioiden soveltamista yhdessä, mikä vaikuttaisi olevan varsin yleinen käytäntö esimerkiksi neurotieteissä (Craver 2007, 157-8).

Esimerkiksi sitä, onko hypothalamuksen spatiaalisten karttojen muodostamiskapasiteetti relevantti hiiren suunnistamiselle Morrisin vesisokkelossa, voitaisiin testata niin, että yhtäältä haetaan evidenssiä häiritsemällä hypothalamuksen toimintaa ja toisaalta stimuloimalla sen toimintaa. Mikäli molemmat interventioasetelmat puoltavat relevanttiuden puolesta, silloin voidaan sanoa, että jokin hypothalamuksen toiminto on tukevapohjaisesti havaittu olevan muistimekanismin komponentti.

Tosin vielä on mahdollista esittää kolmas interventioasetelma, joka mahdollistaa entistä tukevapohjaisemman evidenssin. Kyse on nimittäin niin sanotuista top-down-interventioista, joissa manipulatiivinen vaikutus kohdistetaan koko mekanismiin (Craver & Bechtel 2007; Craver 2007, §8.3). Mainitun kaltaista interventiota nimitetään yleensä ”aktivointi-interventioksi”, sillä sen on tarkoitus ”aktivoida” ylhäältä alaspäin

mekanismin osat (Craver 2007, 151). Koehenkilö saatetaan esimerkiksi asettaa tekemään jotakin kognitiivista tehtävää samalla, kun tämän aivojen metabolista aktiivisuustasoa (verenkiertoa) kuvannetaan jollakin laitteistolla (tyypillisesti fMRI- tai PET-kuvantaminen). Esimerkkinä voisi olla muun muassa emootioiden tunnistamiskoe, jossa koehenkilön asetetaan katselemaan toisen ihmisten ilmeitä koenäytöltä. Jos ilmenisi, että mantelitumakkeen metabolinen aktiivisuustaso nousee aina vihaisten kasvojen ilmaantuessa koenäytölle, tällöin voitaisiin päätellä, että jokin mantelitumakkeen prosessi on mahdollisesti relevantti vihan tai aggression tunnistamisen kannalta.

Myös top-down-interventioilla on omat haasteensa. Keskeisempänä näistä on *korrelaatio*-haaste (Craver 2007, 151-2). Aivokuvantaminen osoittaa parhaimmillaan sen, että jonkin aivoalueen aktiivisuus korreloi jonkin kognitiivisen prosessin kanssa. Aivokuvantaminen voidaan toisin sanoen analogisesti yhtäläistää siihen, että joku havaitsee barometrin elohopean laskevan ja tämän havainnon valossa päättelee, että myrsky on tuloillaan. Aivokuvantaminen voi tuottaa informaatiota siitä, missä verenkierto kohenee, mutta tämä verenkierron kohentuminen ei ole se tapahtuman alainen entiteetti, josta ollaan kiinnostuneita esimerkiksi tarkastellessa mantelitumakkeen roolia emootioiden tunnistamismekanismin kannalta.

Toisen haasteen top-down-interventioille muodostaa komponenttien mahdollinen *toonisuus* (Craver 2007, 152; vrt. Craver & Darden 2013, 189). Kyse on siitä, että kaikki relevantit komponentit eivät aina ole aktivoitavissa sikäli kuin ne saattavat olla staattisia tai inhibitorisia. Toisin sanoen aktivointi-interventio tuottaa informaatiota siitä, *missä jotakin tapahtuu*, mutta joskus selityksellisesti relevanttia on myös se, *missä ei tapahdu*. Viimeiseksi mainittu huomio on varsin oleellinen inhibitoristen seikkojen suhteen; voi olla esimerkiksi niin, että jollakin aivoalueella tapahtuu jotakin $[X(\Phi_2)]$, koska jollakin toisella aivoalueella jotakin estetään tapahtumasta jotakin $[X(\Phi_1)]$ (vrt. Craver 2007, 60, 152). Pulma onkin siinä, että aktivointikoe ilmentää $X(\Phi_2)$ -komponentin roolin, mutta on täysin hiljaa sen suhteen, että myös $X(\Phi_1)$ -komponentilla on oma roolinsa ilmiön kannalta, jolloin viimeiseksi mainitun komponentit relevanttius jää vaille huomiota.

Voitaisiin näin ollen ehdottaa, että jotta mekanismia konstituoivat toiminnalliset komponentit voitaisiin identifioida luetettavasti, tarvitaan relevanssiehto, joka selviää

kaikista edellä mainituista haasteista. Craverin konstitutionaalinen relevanssiehto pyrkii ottamaan tämän haasteen vastaan niin, että sen mukaan jokin komponenttimuuttuja $X(\Phi_i)$ on ilmiömuuttujan $S(\Psi)$ kannalta relevantti (eron tekävä) jos ja vain jos (vrt. Craver 2007, 159):

(MM-defi):

- i) Komponenttimuuttuja $X(\Phi_i)$ on ilmiötason muuttujan $S(\Psi)$ aika-avaruudellinen osa¹³,
- ii) jos ideaalisen intervention I_Φ vaikutuksesta komponenttimuuttuja ottaa arvon $X(\Phi_i) = \phi_1$ ja ilmiötason muuttuja $S(\Psi)$ ottaa arvon sen funktiona, niin että $S(\Psi) = f(\phi_1)$, ja
- iii) jos ideaalisen intervention I_Ψ vaikutuksesta ilmiö tason muuttuja ottaa arvon $S(\Psi) = \psi_1$, silloin komponenttimuuttuja ottaa arvon sen funktiona niin, että $X(\Phi_i) = f(\psi_1)$.

MM ehdon nojalla jokin seikka on siis mekanismin komponentti, jos ja vain jos annetun komponenttiehdokkaan manipuloiminen tuottaa eroja mekanismissa, ja manipulaatio kohdistettuna mekanismiin tuottaa eroja komponenttiehdokkaassa (Craver 2007, 141, 159). Käytännössä MM-ehto vaatii sitä, että alatason komponenttimuuttujan $X(\Phi_i)$ ja ilmiötason $S(\Psi)$ -muuttujan välillä täytyy vallita eron tekävä suhde, mikäli $X(\Phi_i)$ on selityksellisesti relevantti $S(\Psi)$ -muuttujan suhteen. MM-ehto edellyttää tosin sitä, että eron tekävän suhteen pitää olla symmetrinen: $S(\Psi)$ -muuttujan täytyy olla $X(\Phi_i)$ -muuttujan eron tekijä ja $X(\Phi_i)$ -muuttujan täytyy olla $S(\Psi)$ -muuttujan eron tekijä. Tämän symmetrisuuden ongelmallisuutta tullaan havainnollistamaan myöhemmin. Tosin tässä vaiheessa voidaan todeta, että tieteen tekemisen näkökulmasta kyseessä näyttäisi olevan hyvin käytännöllinen relevanssiehto, sillä tämä symmetrinen eron tekävä suhde näyttäisi olevan jotakin, minkä testaamiseen suurin osa tieteellisistä kokeista pyrkii. Tosin sanoen tieteelle tutuissa interventiokekeissa juuri testaan sitä, onko suhde havaittavissa stimulaatio-, häirintä- ja aktivointi-interventioiden suhteen. Yksinkertaisesti MM-ehdon vaikuttavuus näyttäisi nousevan siitä, että se on tieteenfilosofisena muotoiluna

¹³ Tässä on syytä todeta, ettei Craver alkuperäisessä muotoilussaan ole eksplisiittinen sen suhteen, mitä hän tarkoittaa osarelaatiolla (parthood). Esimerkiksi Bert Leuridan (2012) syyttää Craveriä tämän kaltaisesta huolimattomuudesta (ks. myös Kaiser & Krickel 2016, 753). Usein kirjallisuudessa Craverin tarkoittama osarelaatio kuitenkin tulkitaan niin, että se sisältää sekä ajallisen että avaruudellisen osarelaation. Tämä sama tulkinta tehdään myös tässä yhteydessä.

harvinaisen yhteensopiva sen kanssa, kuinka tiedettä tehdään useilla aloilla ja varsinkin neurobiologisilla aloilla.

Esitän seuraavaksi, kuinka MM-ehdon nojalla voitaisiin mahdollisesti esittää ratkaisuehdotus sen suhteen, miksi visuaalisen aivokuoren (VI) aktivoituminen on osa näkömekanismia ja miksi sydämen toiminta ei ole. Tarkastellaan ensin bottom-up interventioita. Ensinnäkin voidaan oletetusti todeta, että sydämen stimulointi ei tuota näkömekanismia. Toisaalta voidaan olettaa, että vakavan asteen häirintäinterventio (esimerkiksi sydänkohtaus) saattaa tuottaa muutoksia näkömekanismeissa. Tosin tämä yhteys ei riitä osoittamaan relevanttiutta, koska sydämen toiminta ei muutu top-down-interventioiden suhteen, eli silloin kun manipuloidaan näkömekanismia. Kuvitellaan esimerkiksi asetelma, jossa koehenkilö suorittaa visuaalisia tehtäviä. Visuaalisten tehtävien suorittaminen on interventio, jonka johdosta näkömekanismeille relevantilla VI-alueella ilmenee jonkin asteista aktivoitumista, mutta sen johdosta tuskin ilmenee muutoksia sydämen toiminnassa. Lisäksi stimulointi ja häirintätyypin interventiot VI:n suhteen tuottavat oletetusti muutoksia näkömekanismeissa. Toisin sanoen MM-ehdon nojalla VI aktivoituminen on näkömekanismin konstitutionaalinen komponentti, koska se on eron tekijä mekanismin kannalta sekä bottom-up -interventioiden että top-down-interventioiden nojalla.

Todetun perusteella vastaus siihen, kuinka interventionismin soveltaminen mekanistisen selittämisen teoriaan käy yksiin (S5) ehdon kanssa, voisi olla seuraava. Komponentti $X(\Phi_i)$ on konstitutionaalinen komponentti, jos ja vain jos se tyydyttää MM-ehdon.

Ehdotan seuraavaa muodollisempaa päättelyä tämän ilmaisemiseksi:

- K) Mekanistisesti selitettävän ilmiön konstitutionaalinen selitys identifioi mekanismin toiminnan mahdollistavat kausaaliset komponentit.
 - Rel) Komponentti on konstitutionaalisesti relevantti mekanistisen kokonaisuuden toiminnalle, jos se tyydyttää mutuaalisen manipulaatioehdon.
 - MM) Komponentti tyydyttää mutuaalisen manipulaatioehdon, jos ja vain jos komponentti on mekanismin aika-avaruudellinen osa, sitä manipuloimalla tuotetaan muutoksia ilmiössä ja ilmiötä kokonaisuutena manipuloimalla tuotetaan muutoksia komponentissa.
- 1) $X(\Phi_i)$ on $S(\Psi)$ aika-avaruudellinen osa
 - 2) Jos ideaalisen intervention vaikutuksesta komponenttimuuttuja ottaa arvon $X(\Phi_i) = \varphi_i$, silloin ilmiömuuttuja ottaa arvon sen funktiona, niin että $S(\Psi) = f(\varphi_i)$

- 3) Jos ideaalisen intervention vaikutukset ilmiömuuttuja ottaa arvon $S(\Psi) = \psi_i$, silloin komponenttimuuttuja ottaa arvon sen funktiona, niin että $X(\Phi_i) = f(\psi_i)$,
- 4) $X(\Phi_i)$ tyydyttää mutuaalisen manipulaatio ehdon \parallel MM, 1, 2
- 5) $X(\Phi_i)$ on relevantti ilmiön $S(\Psi)$ kannalta \parallel Rel
- 6) $X(\Phi_i)$ on relevantti $S(\Psi)$:n konstitutionaalisen selitykselle. \parallel K

Toisaalta MM-ehto herättää monia kysymyksiä siitä, mitä top-down- ja bottom-up-interventiot ovat. Näyttäisi siltä, että mekanistit eivät ole valmiita hyväksymään näiden interventioiden samaistamista kausaalisuuden kanssa (Craver & Bechtel 2007; Craver 2007, 195). Tämä johtuu siitä, että top-down-interventioita ei ole mekanistisessa kannassa haluttu yhdistää mysteeriseen top-down-kausalisuuteen. Sanalla sanoen halutaan kieltää se, että ylemmän tason syillä voisi olla alemman tason seurauksia. Craverin mukaan tämä on seurausta siitä yleisesti hyväksytystä kausalisuuteen liittyvästä huomiosta, että syiden täytyy edeltää ajallisesti niiden seurauksia (Craver 2007, 153). Tähän voidaan lisätä se huomio, että syy-ja-seuraus-suhde voi vallita vain ontologisesti erillisten seikkojen välillä. Tasojen välinen top-down-kausatio on siis tässä mielessä ongelmallinen sikäli kuin ylemmän tason seikkoja ei ymmärretä vahvassa mielessä emmergenteiksi (ks. osio 3).

Tämä ongelmallisuus ilmenee parhaiten ”kausaalisen poissulku”-argumentin (*causal exclusion*) kautta, jonka parhaiten lienee muotoillut Jaegwon Kim (1998, 37-38). Kimin mielenfilosofinen argumentti koskee alkuperäisessä muotoilussaan mentaalista kausaatiota, mutta sitä voidaan soveltaa tässä yhteydessä seuraavalla tavalla. Kuvitellaan, että on olemassa seikka M, joka on ylemmällä konstitutionaalisella tasolla kuin P ja jonka uskotaan olevan jonkin toisen seikan P* kausaalinen (top-down) syy. Kuvitellaan lisäksi, että seikka P on yksin tai itsessään riittävä syy seikalle P* ja että nämä molemmat ovat samalla konstitutionaalisella tasolla. Konstitutionaalisuuden nojalla voidaan olettaa, että koska P on M:n aika-avaruudellinen osa, niin silloin M:n suhteen ei voi ilmetä eroja ilman, että eroja ilmenee samalla P:ssä. Nyt voidaan kysyä, mikä kausaalinen rooli M:llä on P* suhteen. Jos M ei tuota eroja P*:ssä ilman, että tämä eron tekevyys välittyy P kautta (so. kausaalisen polun $P \rightarrow P^*$ kautta), silloin M näyttää olevan kausaalisesti täysin redundantti P* suhteen. Sanalla sanoen top-down-vaikutus M:n ja P*:n välillä on joko mahdoton tai sitten täysi mysteeri.

Poissulkuargumentti on haaste mutuaaliselle manipulaatiolle. Yhtäältä sen takia, että konstitutionaaliset osat ovat ylemmän tason avaruudellisia osia. Toisaalta sen takia, että

Craverin mukaan ylätason ja alatason tapahtumat ovat ajallisesti synkronoituja (2007, §4.2.2.). Tässä kannattaa muistaa se, mitä todettiin osiossa 3 heikosta emergenssistä. $X(\Phi_1)$ ja $S(\Psi)$ eivät voi kausaalisesti vaikuttaa toisiinsa koska ne eivät ole varsinaisesti ontologisesti toisistaan erillisiä objekteja (vrt. Craver 2007, 153). Tätä voidaan havainnollista esimerkillä, jossa X alkaa pelata tennistä ja samalla X:n verensokeriyritykset alkavat laskea (ks. Craver & Bechtel 559-560). Sitä, että X alkaa pelata tennistä, voidaan pitää top-down-interventiona. Se, mitä ei haluta sanoa, on se, että tämä interventio *aiheuttaa kausaalisesti* X:n verensokeriyrityksien laskemisen, koska nämä kaksi asiaa (tenniksen pelaaminen ja verensokeriyrityksien laskeminen) tapahtuvat samanaikaisesti ja saman avaruudellisen objektin suhteen. X:n glukoosimetabolismissa syntyvät muutokset ovat aina solutason vaikutuksia, jotka ovat jonkin solutason (ei makrotason) syyn seurauksia. Tietenkin tiettyssä mielessä vastaus kysymykseen ”miksi X:n verensokeriyritykset laskivat” on se, että X alkoi pelata tennistä. Tosin sanotun perusteella voidaan päätellä, ettei tämä ”selitys” näytä jäljittävän kausaalisuutta. Tulkiten Craverin ja William Bechtelin näkemystä niin, että kyseessä näyttäisi olevan tällöin ei-kausaalinen eli *konstitutionaalinen eron tekvä suhde*, jonka makrotason interventio jäljittää. Tästä voidaan johtaa seuraava kiteyttävä ehto:

EK: Tasojenvälinen riippuvaisuus ei ole kausaalista vaan konstitutionaalista riippuvaisuutta.

Craverin ja Bechtel kannassa top-down-kausalisuus yritetään siis ”selittää pois” alatason välittämien vaikutuksien avulla (Craver & Bechtel 2007, 547).

Konstitutionaalinen relevanssi halutaan säilyttää kaksisuuntaisena relaationa ja kausaalinen relevanssi yksisuuntaisena. Väittäisin, että nyt ei ole vaikeaa nähdä, että jossain on menty pieleen ja pahasti. Juuri todettu näyttäisi olevan ristiriidassa sen kanssa mitä edellä on todettu eron tehvyydestä ja interventioista. Käytännössä EK- ja MM-ehdot yhdessä johtavat vakaviin ongelmiin. Osoitan tämän seuraavaksi tarkemmin.

6.2. Mutuaalisen manipulaation dilemma

MM-ehdon ongelmat nousevat juurikin siitä, että MM-ehdon intuitiivinen vaikuttavuus ostetaan loogisen selkeyden hinnalla. Näyttäisi siis siltä, että Craver on MM-ehdossaan sekoittanut keskenään episteemisen ja ontologisen ehdon (Glennan 2017, 44). Se näyttäisi tarjoavan hyvän episteemisen ehdon, miten mekanismin toiminnalliset osat erotetaan ei-toiminnallisista. Tästä episteemisestä ehdosta ei kuitenkaan voida

yksiselitteisesti hypätä konstitutionaalisuutta koskeviin ontologisiin johtopäätöksiin ongelmitta. Tämän sekaannuksen vuoksi MM-ehto on ollut kritiikin kohteena (Leuridan 2012; Romero 2015; Baumgartner and Gebharder 2016). Tätä kritiikkiä on tarkoitus käydä läpi seikkaperäisemmin.

Keskeinen argumentti on ollut se, että konstitutionaalinen relevanssi, MM-ehdon mukaisesti muotoiltuna woodwardilaisen systeemin laajennuksena, on luonteeltaan kausaalinen relevanssiehto, joka on puolestaan kestävä johtopäätös (Leuridan 2012, §5; Romero 2015, §2.3). Toisin sanoen se on ristiriidassa EK-ehdon kanssa.

Argumentin ytimessä on näkemys, että mutuaalinen manipulatiivisuus edellyttää kaksisuuntaista kausaalisuutta, mikäli se muotoillaan Woodwardin esittämän IV-ehdon mukaisesti. Tämä johtuu siitä, että MM perustuu ongelmalliselle poikkileikkaavan intervention käsitteelle, jossa oletetaan ilmiön ja sen osien välillä vallitsevan eron tekevä suhde. Tätä ongelmallisuutta voidaan aloittaa ilmentämään esittämällä tasot poikkileikkaavan interventio muodollisesti:

PK: Jokaisen muuttujan $X(\Phi_i)$ suhteen on olemassa sellainen $S(\Psi)$ -interventio I_Ψ niin, että I_Ψ tuottaa arvomuutoksen sekä muuttujassa $S(\Psi)$ että muuttujassa $X(\Phi_i)$ yhden ja saman kausaalisen reitin välityksellä, ja muuttujan $S(\Psi)$ suhteen on olemassa jokin sellainen $X(\Phi_i)$ -interventio I_Φ niin, että I_Φ tuottaa arvomuutoksen sekä muuttujassa $S(\Psi)$ että muuttujassa $X(\Phi_i)$ *yhden ja saman kausaalisen reitin välityksellä*.

Tästä seuraa IV-I i-ehdon nojalla se, että interventio täytyy olla aina sekä $S(\Psi)$:n että $X(\Phi_i)$:n (manipulatiivinen) syy. Tosin ongelmaksi muodostuu se, että interventio voidaan tässä tulkita kolmella eri tavalla (Baumgartner & Gebharder 2016, 742). Tätä monimielisyyttä voidaan ilmentää kolmella kausaalisella kaaviolla:

- A) $I_\Psi \rightarrow S(\Psi) \rightarrow X(\Phi_i)$ [top-down-interventio]
- B) $I_\Phi \rightarrow X(\Phi_i) \rightarrow S(\Psi)$ [bottom-up-interventio]
- C) $X(\Phi_i) \leftarrow I_\Psi \rightarrow S(\Psi)$ [paksukätinen interventio]

PK-ehdon mukaan vain (A) ja (B) ovat sallittuja vaihtoehtoja, koska ne ilmentävät poikkileikkaavaa interventiota. Tosin edellä esitetystä MM-ehdon edellyttämästä EK-ehdosta seuraa se, että vaihtoehdot (A) ja (B) ovat poissuljettuja vaihtoehtoja, koska niiden nojalla konstitutionaalinen relevanssi olisi kausaalinen riippuvuussuhde. Toisin

sanoen jäljelle jää vaihtoehto (C), jossa interventio aiheuttaa muutokset $S(\Psi)$ ja $X(\Phi_i)$ muuttujissa kahden toisistaan erillisen kausaalisen reitin kautta. Tämä on ristiriidassa PK-ehdon kanssa, eikä se ole yhteensopiva alkuperäisen interventionismin kanssa, koska se rikkoo osiossa 5 esitetyn IV-I (ii)-ehtoa. Todetun perusteella MM-ehtoa ei voida muotoilla alkuperäisen interventionismin laajennuksena, mikäli EK- ja PK-ehdoista molemmista halutaan pitää kiinni.

Seuraukset tästä yhteensopimattomuudesta ovat absurdit, sillä sen mukaan top-down- ja bottom-up-interventiot eivät olisi mahdollisia. Näissä interventioissa havaittaisiin yksinkertaisesti vain ylätasoon ja alatasoon välinen kovarianssi. Ψ -interventio olisi tässä tapauksessa sekä $S(\Psi)$ että $X(\Phi)$ yhteinen manipulatiivinen syy, eikä sitä voisi siten kutsua varsinaiseksi top-down-interventioksi. Kyse olisi tällöin niin sanotusta paksukätisestä interventiosta (Gebharter & Baumgartner 2016; vrt. Scheines 2005, 932). Muodollisemmin ilmaistuna paksukätisyys voidaan muotoilla seuraavasti (vaihtoehtoinen muotoilu ks. Baumgartner & Gebharter 2016, 752):

Paksu-I-def: Joukon V komponentit $\{\Phi_1, \Phi_2, \dots, \Phi_n\}$ ja ilmiömuuttuja $S(\Psi)$ tyydyttävät paksukätisyyden, jos ja vain jos jokainen $S(\Psi)$ -muuttujaan kohdistettu interventio I_Ψ on sekä $S(\Psi)$ -muuttujan että jonkun $X(\Phi_i)$ -muuttujan syy, niin että $X(\Phi_i) \in V$ mutta kuitenkin niin, että I_Ψ välittää kausaalisen vaikutuksena $S(\Psi)$ -muuttujan että $X(\Phi_i)$ -muuttujan suhteen *kahden erillisen kausaalisen polun kautta*.

Paksukätinen interventio tuottaa siis muutoksia komponentissa sellaisen kausaalisen reitin kautta, joka ei kulje fenomenaalisen $S(\Psi)$ -muuttujan kautta. Seurauksena tästä kaikesta on se, että ei voida suorittaa yhtäkään sellaista interventiota $S(\Psi)$ suhteen, joka sinänsä tuottaisi arvomuutoksia $X(\Phi)$ -muuttujassa $S(\Psi)$:n arvomuutoksien funktiona. Näinmuodoin ollaan tultu pisteeseen, jossa Craverin MM-ehto kohtaa dilemman:

D-1: Jos mutuaalinen manipulatiivisuus määritellään PK-ehdon mukaisesti, on konstitutionaalinen suhde kaksisuuntainen kausaalinen riippuvuussuhde.

D-2: Jos mutuaalinen manipulatiivisuus määritellään EK-ehdon mukaisesti mutta ilman PK-ehtoa, silloin konstitutionaaliset tasojen välillä ei vallita eron tekevää suhdetta.

D-1 on ongelmallinen, koska siinä edellytetään filosofisesti kyseenalaista kaksisuuntaista kausaalisuutta (Leuridan 2012). Tosin myös D-2 on ongelmallinen, koska se vie pohjan poikkileikkaavilta interventioilta (Baumgartner & Gebharter 2016; Romero 2015), jolloin relevanssiehto uhkaa kaatua; relevanssihan perustuu selityksellisten seikkojen väliseen eron tekävään suhteeseen.

Tuoreessa kirjallisuudessa on esiintynyt kolme tapaa, jolla kohdata dilemman johtopäätös: a) heikennetään mutuaalisen manipulaatiokriteerejä (Baumgartner & Casini 2017); b) hyväksytään tasojenvälinen kausaalisuus (Krickel 2017; 2018); c) hylätään mutuaalinen manipuloitavuuskriteeri (Leuridan 2012) tai jätetään huomiotta (Gebharter 2016). Viimeiseksi mainittu ratkaisuehdotus jätetään tässä yhteydessä käsittelemättä, sillä MM-ehdon hylkääminen sanotun perusteella on hätäinen johtopäätös. MM-ehdon pelastamisen voidaan katsoa olevan yrittämisen arvoista sikäli kuin sen lupaukset ovat suuret. Ratkaisua on siis haettava hyväksymällä joko D-1- tai D-2 johtopäätös. Tarkoitukseni on siis osoittaa, ettei MM-ehdon pelastaminen ole täysin mahdotonta, mutta että nykyisissä muodoissaan tarjotut ratkaisuehdotukset ovat riittämättömiä.

6.2.1. Paksukätinen ratkaisu

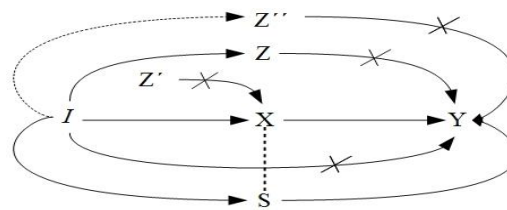
Michael Baumgartner ja Lorenzo Casini (2017) päätyvät siihen johtopäätökseen, että PK on hylättävä, sillä heidän mukaansa Craverin MM-ehtojen taustalla oleva perusajatus on virheellinen. Heidän mukaansa ilmiön ja sen osien välillä sinänsä ei voi vallita eron tekävää suhdetta (2017, 216). Toisin sanoen tässä ratkaisumallissa, vastausta dilemmaan yritetään löytää D-2 polun kautta. Käytännössä Baumgartner ja Casini hyväksyvät paksukätiset interventiot eli sen, että mekanismin ilmiötasolla ja sitä konstituivalla alatasolla eroja tuottava interventio tuottaa erot kahden erillisen kausaalisuuden välityksellä. Tämä voidaan asettaa heidän mukaansa yksiin Woodwardin uuden, lyhennetyn interventionistisen kannan kanssa (Woodward 2015; Baumgartner & Casini 2017, 221-2). Aiemmassa teoriassaan Woodward salli vain yhden muuttujan interventiot (2003, 98), kun taas uudemmassa teoriassaan (2015, 333–334) hän luo pohjaa interventiokäsitykselle, jossa interventio voi kohdistua useisiin muuttujiin samanaikaisesti, olettaen että nämä muuttujat ovat jonkinlaisessa supervenienssisuhteessa. Näin ollen interventiot voivat mukauttaa samanaikaisesti sekä makrotason muuttujaa että supervenienssipohjan muuttujaa kahden eri kausaalisuuden välityksellä, jolloin paksukätiset interventiot voidaan sallia sillä hinnalla, että interventiot eivät ole koskaan poikkileikkaavia. Kutsuttakoon näitä *sallituiksi*

paksukätisiksi interventioiksi. Näin muodoin voidaan ehdottaa seuraavaa modifikaatiota edellä esitettyihin IV-I-ehdoin (vrt. Baumgartner & Gebharder 2016, 745):

I-ii*) Kaikki kausaaliset reitit muuttujasta I muuttujaan Y, kulkevat muuttujan X kautta tai jonkin sellaisen muuttujan S kautta, joka supervenisoii muuttujaa X.

I-iii*) I ei korreloi minkään sellaisen muuttujan (Z'') kanssa, joka on kausaalisesti relevantti muuttujan Y suhteen niin, että se ei ole sellaisella kausaalisella polulla, joka kulkee muuttujan X kautta; tai sellaisella kausaalisella polulla, joka kulkee sellaisen muuttujan S kautta, joka supervenisoii muuttujaa X.

Näitä ehtoja voidaan myös havainnollistaa seuraavalla kaaviolla (vrt. Krickel 2018, 61):



Kaavio 3

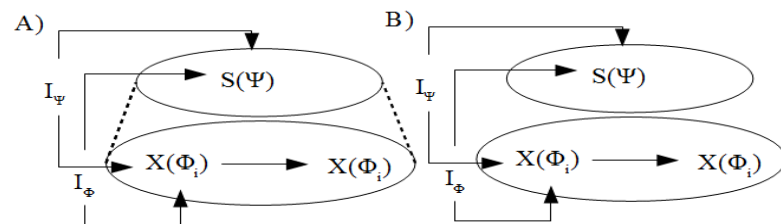
Kaavio 3 osoittaa sen, että intervention vaikutus Y suhteen voi kulkea jonkin sellaisen S-muuttujan kautta, joka supervenisoii X-muuttujaa. Tällöin sallittu paksukätinen interventio on sellainen, jonka paralleelit seuraukset (X ja S) ovat supervenienssisuhteessa. Näin ollen voidaan ehdottaa, että komponentti $X(\Phi_i)$ konstituoii ilmiötä $S(\Psi)$ jos ja vain jos (vrt. Baumgartner & Wilutzky 2017, 1115):

MM-def₂:

- 1) $X(\Phi_i)$ on $S(\Psi)$ aika-avaruudellinen osa;
- 2) Muuttujalle $X(\Phi_i)$ on olemassa sellainen *sallittu paksukätinen interventio* I_Φ muuttujan $S(\Psi)$ suhteen, jonka vaikutuksesta $X(\Phi_i)$ ottaa arvon ϕ_1 ja $S(\Psi)$ arvo muuttuu samanaikaisesti;
- 3) Muuttujalle $S(\Psi)$ on olemassa sellainen *sallittu paksukätinen interventio* I_Ψ jonkin $X(\Phi_i)$ suhteen, jonka vaikutuksesta $S(\Psi)$ ottaa arvon ψ_1 ja $X(\Phi_i)$ arvo muuttuu samanaikaisesti.

Tämä ei kuitenkaan ole riittäväehto, sillä paksukätisistä interventioista ei voi päätellä suoraan konstitutionaalisuuteen (Baumgartner & Casini 2017, 221-5). Kyse on siitä, että *MM-def₂*-ehdolla ei ole niitä samoja hyveitä kuin alkuperäisellä *MM-def₁*-ehdolla, joka perustuu poikkileikkaaviin interventioihin. Uuden määritelmän ei voida siis katsoa tarjoavan metodia, jolla konstitutionaalisuus voitaisiin identifioida. Nimittäin koska

paksukätiset interventiot eivät voi olla poikkileikkaavia, eivät ne voi tuota suoraa informaatiota konstitutionaalisuudesta, jolloin niiden tuottama informaatio ei paljasta konstitutionaalista suhdetta. Lopputulos on se, että paksukätinen interventionismi sallii konstitutionaalisuuden alideterminoitumisen (Baumgartner & Casini 2017, 221-5). Esimerkiksi seuraavat kaksi mallia ovat paksukätisen interventioiden nojalla empiirisesti erottamattomia (vrt. Baumgartner & Casini 2017, 223):



Kaavio 4

Toisin sanoen paksukätisen interventiosta saadaan dataa, joka sinänsä ei kerro, mikä relaatio vallitsee intervention kohteina olevien muuttujien välillä. Tämä synnyttää alideterminaatio ongelman: paksukätinen intervention pohjalta voidaan yhtäältä puolttaa kahden vastakkaisen mallin puolesta; mallissa A on kyse konstitutionaalisesta relaatiosta, kun taas mallissa B ei ole. Tämä on seurausta siitä, että paksukätinen manipulaatio antaa vain informaatiota siitä, että ylätasen ja alatason muuttujat ovat yhteisen syyn linkittämiä, mutta tämä ei ole informaatiota konstitutionaalisuudesta. Baumgartnerin ja Casinin mukaan tämän kaltainen manipuloitavuus ei ole siis jotakin, mistä konstitutionaalisuusrelaatio voitaisiin päätellä *suorasti* (ibid.). Tarkoitukseni on seuraavaksi laatia esimerkki tämän kaltaisesta alideterminaatiosta.

Kuvitellaan, että koehenkilön S muistimekanismissa ilmenee muutoksia $K(\Psi)$, kun S:lle annetaan lääkettä L (interventio). Kuvitellaan lisäksi, että hallussa on ”paksukätistä” evidenssiä, jonka mukaan L:n vaikutuksesta S:n kehossa aina tapahtuu muutoksia jonkin tapahtuman alaisen entiteetin $Y(\Phi)$ suhteen. Tällöin ollaan alideterminaatio skenaariossa: lääkeintervention nojalla ei voida yksiselitteisesti päätellä, onko $Y(\Phi)$ ilmiötä $K(\Psi)$ konstituiva vai ei, sillä $Y(\Phi)$ voi olla pelkkä lääkkeen aiheuttama sivuvaikutus, jolloin $Y(\Phi)$:llä ei olisi mitään osaa muistimekanismin suhteen (kaavion skenaario b). *Käytännössä tämä ei ole ongelmallinen tilanne*, koska tutkimalla $Y(\Phi)$ ja $K(\Psi)$ välistä linkittyneisyyttä eri asetelmista käsin (esim. eri lääkkeillä tai kontrolloimalla muita muuttujia) voidaan saada tukevapohjaista evidenssiä, jonka nojalla tämä sivuvaikutteisuus voidaan hyvin perustein joko todentaa hylätä. Ongelma

koskee sitä, että mikään rajallinen interventioista johdettava testikokonaisuus ei ole niin ideaalinen, että siitä voitaisiin *suoraan* tehdä ontologisia johtopäätöksiä $Y(\Phi)$:n konstitutionaalisesta roolista. Toisin sanoen yhdenkään paksukätisen intervention tai niiden kokonaisuuden nojalla sinänsä ei voida päätellä, ollaanko kaavion 4 mukaisesti skenaariossa (a) vai (b). Baumgartner ja Casini näyttäisivät esittävän, että tämä kuilu skenaarioiden välillä on mahdollista ylittää vain abduktiivisesti, eli ei interventioiden nojalla *per se* (Baumgartner & Casini 2017, 225-32). Hypoteettinen konstitutionaalinen suhde $Y(\Phi)$:n ja $K(\Psi)$:n välillä on *vain* paras mahdollinen selitys sille, miksi $Y(\Phi)$:n on havaittu monista eri koeasetelmista käsin systemaattisesti ilmenevän yhdessä $K(\Psi)$:n kanssa.

Käytännössä tämä tarkoittaa sitä, että yhteiselle syyille perustuvan linkittyneisyyden murtumattomuutta testataan mahdollisimman laajoilla interventiokokonaisuuksilla ja mahdollisimman laajojen taustaehtojen valossa. Murtumattomuudesta voitaisiin puolestaan päätellä abduktiivisesti, että jokin komponentti $X(\Phi_i)$ konstituoii mekanismia $S(\Psi)$. Jonkin komponenttien joukon $V_i = \{X(\Phi_1), \dots, X(\Phi_n)\}$ voidaan abduktiivisesti päätellä konstituoivan jotakin ilmiötä $S(\Psi)$, kun sekä V_i :n että $S(\Psi)$:n voidaan *systemaattisesti* havaita olevan yhteisen syyn linkittämiä. Kyse on siitä, että paksukätinen manipulaatio voi vain ja ainoastaan empiirisesti falsifioida sen, että jokin $X(\Phi_i)$ -seikka olisi ilmiön $S(\Psi)$ komponentti (Baumgartner & Wilutzky 2017, 1120; Baumgartner & Casini 2017; 231). Näin ollen koska manipuloitavuus sallii vain falsifioinnin, yrittävät Baumgartner ja Casini pelastaa konstitutionaalisen relevanssiehdon esittämällä MM-def₂-ehdon lisäksi abduktiivisen työkalun, jolla konstitutionaalisuus voidaan päätellä datasta, jota ei ole toistaiseksi laajamittaisten manipulaatioiden nojalla voitu falsifioida. Vastaus ei ole kuitenkaan tyydyttävä, mikäli halutaan pitää kiinni alkuperäisen teorian hyveistä.

Ensinnäkin abduktiivisen päättelyn mukaan tuominen johtaa siihen, että yksinkään rajallinen testauskokonaisuus ei voi lopullisesti oikeuttaa konstitutionaalisuuden paljastumista. Tämä näyttäisi olevan johtopäätös, jonka Baumgartner ja Casini ovat valmiit hyväksymään (2017, 13). Väitän kuitenkin, että Baumgartner ja Casini näyttäisivät kuitenkin vähättelevän tämän seurauksia. Isoksi ongelmaksi voidaan väittää muodostuvan nimittäin se, että konstitutionaalisuudesta tulee todetun nojalla löyhä määritelmällinen. Tarkemmin sanoen MM-def₂-ehtojen nojalla konstitutionaalisuus ei ole selkeämääritelmällisellä perustalla. Tämä sama löyhä määritelmällisyys siirtyy

koskemaan myös konstitutionaalista selittämistä. Voidaan siis väittää, että kanta ei ole yhteensopiva ontittisen kannan kanssa sikäli kuin paksukätiset interventiot eivät tuota suoraa evidenssiä maailman konstitutionaalisesta rakenteesta (olettaen, että myös konstitutionaalinen rakenne mielletään ontittiseksi rakenteeksi).

Toiseksi ei ole selvää, miten paksukätisyys on yhteen sovitettavissa tieteellisen käytännön kanssa. Baumgartner ja Casini eivät näytä edes olevan motivoituneita yrittämään muotoilla MM-def₂-ehtoa niin, että se heijastaisi jollakin tavalla tieteellistä käytäntöä. Ei ole toisin sanoen selvää, miten abduktiivisten päätelmien tekeminen paksukätisten interventioden pohjalta vastaa sitä, miten tutkijat päättelivät relevantit komponentit. Craverin teoria näyttäisi olevan lähempänä tieteellistä käytäntöä. Tämä haaste näyttäisi olevan sellainen, ettei siihen edes voida vastata filosofian keinoin. Se, että tekevätkö tieteilijät abduktiivisia päätelmiä, näyttäisi olevan kysymys, jonka suhteen tieteenfilosofian tulisi konsultoida tieteesosiologiaa tai -psykologiaa. Toinen metodisuutta koskeva ongelma on poikkileikkaavien interventioiden puuttuminen. Paksukätisyyden nojalla ei voida tehdä eroa top-down- ja bottom-up-interventioiden välillä, koska paksukätinen interventio ei ole poikkileikkaava. Tieteellinen käytäntö näyttäisi kuitenkin vahvasti tukevan top-down- ja bottom-up-interventioiden kaltaisia menetelmiä. Abduktiivinen vastaus vie siis pohjan PK-ehdolle perustuvalta teorialta, joka on lähempänä tieteellistä käytäntöä. Voidaan siis väittää, että on hyvin epätodennäköistä, että MM-def₂-ehtoa voitaisiin muotoilla koskaan tavalla joka tulee yhtä lähelle tieteellistä käytäntöä kuin sellaiset kannat, jotka sallivat poikkileikkaavat interventiot (kuten MM-def₁ ja seuraavaksi esiteltävä MM-def₃).

Lisäksi tässä yhteydessä voidaan ottaa huomioon myös kolmas ongelma, nimittäin se, että paksukätisyyden nojalla on mahdotonta ottaa huomioon ilmiön järjestäytymistä ajan suhteen (Krickel 2018, 63). Ilmeistä on, että jonkin intervention vaikutuksesta kaikki muutokset mekanismin komponenteissa eivät tapahdu yhdellä ja samalla hetkellä, niin kuin paksukätinen interventionismi antaa ymmärtää. Paksukätinen interventionismi ei ota tätä ilmiön aika-avaruudellista luonnetta huomioon, koska siinä ala- että ylätasen muuttujien täytyy muuttua samanaikaisesti. Tästä päästäänkin aasinsiltana Beate Krickelin (2017; 2018) kantaan, jossa tämän kaltainen ajallisuus otetaan huomioon. Mainittua kantaa voidaan pitää D-1 polun kautta lähestyttävänä ratkaisuehdotuksena.

6.2.2. Kausaalinen ratkaisu

D-1 polkua kulkeva ratkaisumalli edellyttää pieniä metafyyysisiä huomioita. Käytännössä edellä esitettiin jo tämä huomio pähkinänkuoressa: ilmiöt ovat aika-avaruudellisesti jaksottuneita (Krickel 2017; 2018). Paremmin sanoen tämä tarkoittaa sitä, ettei ilmiö ole monoliittinen kokonaisuus, joka voitaisiin ilmaista vain yhden jakamattoman muuttujan varassa. Krickelin keskeinen väite onkin se, että suurin osa mekanistisista teorioista ei ole analysoinut ilmiön käsitettä tarpeeksi syvällisellä metafyyysisellä otteella (Kaiser & Krickel 2017). Perusdualismissa ilmiö on nimittäin aina ilmaistu yksittäisellä $S(\Psi)$ -muuttujalla.

Jotta ratkaisun suhteen päästään etenemään, tulee määritellä uudelleen se, mitä tarkoitetaan ilmiöllä eli sillä, että $S(\Psi)$. Tämä uudelleen määrittely alkaa siitä, että ilmiö on systeemi, joka on itsessään tapahtuman alainen entiteetti, jolla on ajallisia että avaruudellisia osia (Kaiser & Krickel 2017; Krickel 2017; 2018). Tapahtuma on oleellinen termi, koska se sisältää ajatuksen siitä, että jotakin on tekeillä vain ja ainoastaan tietyn määrätyn aikaperiodin suhteen. Sanalla sanoen ilmiö on eräänlainen tapahtumasysteemi, joka koostuu aika-avaruudellisista osista, joita voidaan kutsua EIO-osiksi (entity-involving occurent) (Krickel & Kaiser 2017; Krickel 2017, 461; Krickel 2018, 64). Tämän nojalla voidaan antaa seuraavan kaltainen määritelmä EIO-osista.

Tapahtuman alainen entiteetti E1 on toisen entiteetin E2 ajallinen EIO-osa, jos ja vain jos:

Ajalliset-osat-def:

- “(i) [T]he entity involved in E1 is identical with the entity involved in E2;
- (ii) the activity involved in E1 begins later and ends earlier than the activity involved in E2, or the former begins simultaneously with the latter and ends earlier than the latter, or the former begins later than the latter and ends simultaneously with the latter.” (Krickel 2018, 64.)

Ensin mainittu ehto tuo ilmi E1:n ja E2:n välisen identiteetille perustuvan osarelaation (parthood). Vastaavasti toinen ehto tuo ilmi sen, että osarelaatio ajallinen; E1 sisältämän tapahtuman täytyy tapahtua E2 sisältämän tapahtuman aikana. Vielä tulee kuitenkin määritellä se, mitä tarkoitetaan avaruudellisilla EIO-osilla; eli kuinka ajallisesta

osarelaatiosta itsenäinen avaruudellinen osarelaatio määritellään. Määritelmän mukaan entiteetti E1 on toisen entiteetin E2 avaruudellinen osa, jos ja vain jos

Avaruudelliset osat-def:

- “(i) [T]he entity involved in E1 occupies a proper sub-region of the spatiotemporal region occupied by the entity involved in E2;
- (ii) the activity involved in E1 occurs during the activity involved in E2.” (Krickel 2018, 64.)

Näin määriteltynä ilmiö koostuu aina entiteeteistä jotka ovat EIO-osia. Käytännössä ajalliset ja avaruudelliset EIO-osat ovat ontologisesti toisistaan erillisiä asioita, mikä mahdollistaa sen, että niiden välillä voi periaatteessa olla kausaalisia riippuvuussuhteita. Tässä kannattaa muistaa se, mitä todettiin osiossa 5, että mikäli kausaalisuus ymmärretään kontrafaktuaalisesti, ei syy-ja-seuraus-suhde edellytä välttämättä esimerkiksi fysikaalista vuorovaikutussuhdetta syyn ja seurauksen välillä. Krickelin väite näyttäisi siis olevan loppujen lopuksi se, että jos EIO-osat ovat ontologisesti toisistaan erillisiä, silloin niiden välillä on metafyyysisesti mahdollista olla eron tekevä suhde. Tämä mahdollistaa puolestaan sellaisen konstitutionaalisen relevanssiehdon antamisen, joka mahdollistaa dilemman D-1 johtopäätöksen hyväksymisen sellaisenaan. Tämä voidaan määritellä niin, että $X(\Phi)$ on konstitutionaalisesti relevantti $S(\Psi)$ suhteen, jos ja vain jos:

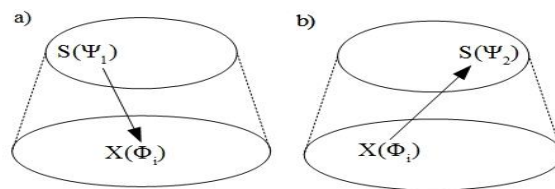
MM-def₃:

- i) $X(\Phi_i)$ on $S(\Psi)$ avaruudellinen EIO-osa
- ii) On olemassa jokin sellainen $S(\Psi)$ ajallinen EIO-osa, joka aiheuttaa kausaalisesti $X(\Phi_i)$
- iii) On olemassa sellainen $S(\Psi)$ ajallinen EIO-osa, joka on kausaalisesti $X(\Phi_i)$ aiheuttama. (Krickel 2018, 65.)

Ehto ii) täyttyy, jos on olemassa sellainen $S(\Psi)$ ajallinen EIO-osa Ψ_1 , jonka suhteen on olemassa ideaalinen interventio I_Ψ niin, että sen vaikutuksesta EIO-osan $\Psi_1 = \psi_1$ arvomuutos johtaa jonkin komponentin $X(\Phi_i) = \phi_i$ arvomuutokseen (Krickel 2018, 64). Vastaavasti ehto iii) täyttyy, jos on olemassa jokin sellainen komponentti $X(\Phi_i)$, jonka suhteen on olemassa ideaalinen interventio I_Φ niin, että sen vaikutuksesta komponentin

$X(\Phi_i) = \varphi_i$ arvomuutos johtaa arvomuutokseen jossakin ilmiömuuttujan $S(\Psi)$ ajallisessa EIO-osassa Ψ_2 (ibid., 65).

Oleellinen johtopäätös on siis se, että ehtojen ii) ja iii) varaan voidaan rakentaa manipulatiivinen systeemi, joka on yhteensopiva mutuaalisen manipulaatioehdon kanssa. Käytännössä Krickelin analyysin tarkoituksena on ollut osoittaa, että konstitutionaalinen relevanssi voidaan esittää sellaisen MM avulla, joka perustuu kaksisuuntaiselle kausaalisuudelle. Tämä on mahdollista vain silloin, kun ilmiöt määritellään niin, että ne koostuvat sekä ajallisista että avaruudellisista osista. Tämän metafyyssisen modifikaation ytimessä on se käsitteellinen silmäkääntötempu, jossa tasojen välisen kausaation mahdollistaa se, että avaruudelliset osat (mekanismin komponentit) ovat kaukaalisessa vuorovaikutuksessa ilmiön ajallisten osien kanssa (ilmiötaso). Graafisesti tätä voidaan kuvata seuraavasti (ibid., 65):



Kaavio 5

Jokainen ilmiö on siis tapahtumasysteemi jonkin aikaperiodin t_i - t_n suhteen, mikä tarkoittaa sitä, että jokaisella ilmiöllä on ajallisia osia, jotka voivat olla sekä top-down-että bottom-up-intervention kohteena. Yllä olevan kaavion mukaisesti (b) esittää bottom-up-intervention jonkin $X(\Phi)$, joka tuottaa muutoksia jossakin ilmiön ajallisessa EIO-osassa Ψ_2 . Vastaavasti (a) osoittaa tyypillisen top-down-intervention kohdistettuna johonkin ilmiön ajalliseen EIO-osaan Ψ_1 , joka tuottaa muutoksia jossakin $X(\Phi_i)$.

Mutuaalisen manipulaatio ehtojen täyttymistä voidaan todetun nojalla yrittää ilmentää Krickelin omalla esimerkillä (Krickel 2018, 64-65). Hiiren navigoiminen sokkelossa voi olla tarkasteltu ilmiö, joka tapahtuu aikaperiodin t_1 - t_4 suhteen. *Ilmiö on siis tapahtuman alainen entiteetti* (hiiri = entiteetti ja hiiren suunnistaminen = aktiviteetti). Näin ollen ilmiö voidaan ”hajottaa” esimerkiksi neljään ajalliseen EIO-osaan (Ψ_1 - Ψ_4). Ilmiön avaruudellisia EIO-osia ovat muun muassa hiiren lihaksien supistuminen tai sen hippokampuksen aktiivisuus, joka vastaa spatiaalisten karttojen muodostumisesta. Ilmiön ajalliset EIO-osat koostuvat puolestaan ajallisista tapahtumista, kuten esimerkiksi siitä, että hiiri asetetaan hetkellä t_1 sokkeloon tai hiiri kääntyy hetkellä t_2 oikealle.

Oleellista on huomata, että ajalliset osat saattavat olla esimerkiksi hetkellä t_3 avaruudellisen tapahtuman syy; esimerkiksi niin, että hiiren hippokampus aktivoituu siinä-ja-siinä määrin hetkellä t_3 , koska hetkellä t_1 ja t_2 tapahtui sitä-ja-sitä. Tämä hippokampuksen aktivoituminen voi olla puolestaan sen ajallisen osan Ψ_4 syy, että hiiri löytää ulos sokkelosta. Kuvitellaan nyt koeasetelma, jossa hiiri asetetaan suunnistamaan sokkelossa ja kokeen suorittavat tutkijat monitoroivat hiiren hippokampuksen aktiivisuutta sekä hiiren käyttäytymistä. Halutaan siis tietää, onko hippokampaalinen aktiivisuus hetkellä t_3 konstitutionaalisesti relevantti. Tulee olla ensinnäkin niin, että on olemassa sellainen ilmiön ajallista EIO-osaa, jota manipuloimalla tuotettaisiin arvomuutoksia hippokampuksen aktiivisuustasossa hetkellä t_3 . Toisin sanoen ehdon (ii) tulee täyttyä. Voidaan esimerkiksi kuvitella, että muutos hetkessä t_1 , jona hiiri aloittaa suunnistamisen, tuottaa muutoksia siinä hippokampuksen aktiivisuustasossa, joka havaitaan hetkenä t_3 . Ehto (ii) täyttyy, mikäli interventio t_1 ajalliseen EIO-osaan on eron tekevä hippokampuksen suhteen hetkellä t_3 .

Vielä todennettava, että (iii) ehto täyttyy. Toisin sanoen täytyy olla olemassa jokin sellainen mahdollinen hippokampuksen aktiivisuuteen kohdistettava interventio, jonka vaikutuksesta jokin ilmiön ajallinen EIO-osa olisi ollut toisenlainen. Esimerkiksi jos hetkellä t_3 hippokampukseen kohdistettaisiin interventio, silloin hiiren sokkelosta ulos löytäminen hetkellä t_4 olisi ollut toisenlainen kuin silloin, jos interventiota ei olisi suoritettu (so. hiiri olisi löytänyt tiensä ulos aikaisemmin tai myöhemmin intervention vaikutuksesta). Käytännössä hippokampuksen aktiivisuus tason hetkellä t_3 täytyy siis olla eron tekevä jokin ilmiön ajallisen EIO-osan suhteen. Tutkijat voivat vastaavasti kysyä, miksi hiiren vatsan aktiivisuus hetkellä t_3 ei ole konstitutionaalisesti relevantti sen mekanismin kannalta, joka mahdollistaa hiiren suunnistamisen ulos sokkelosta. On ilmeistä, ettei ehto ii) täyty: jos mihinkään ilmiön ajalliseen osaan kohdistettaisiin interventio, ei vatsan aktiivisuus olisi sen johdosta sinänsä toisenlainen. Tämä jo riittää rajaamaan vatsan aktiivisuuden irrelevanttina seikkana mekanismin ulkopuolelle.

Kaiken tämän raskaan metafyyssisen määrittelyn jälkeen voidaan todeta, että Krickelin kanta välttää osan abduktiivisen kannan ongelmista. Koska Krickelin kanta mahdollistaa suoran päättelyn kausaalisesta relevanssista konstitutionaaliseen relevanssiin, on konstitutionaalisuus selkeästi määriteltävissä toisin kuin abduktiivisessa mallissa, jossa konstitutionaalisia osia ei voida koskaan määrittää selkeästi. Lisäksi

kanta mahdollistaa poikkileikkaavat interventiot, joten se on lähempänä alkuperäistä MM-ehdolle perustuvaa konstitutionaalisuus relevanssikantaa. (Krickel 2018, 66.)

Krickelin kantaa kuitenkin näyttäisi piinaavan sama ongelma kuin paksukätisyydelle perustuvaa abduktiivisuutta: se ei millään tavalla tee ilmeiseksi sitä, että se olisi yhteensopiva tieteellisen käytännön kanssa. Sanalla sanoen Krickelin analyysi tekee ilmiön ajallisista osista keskeisiä interventiokohteita, mutta ei ole mitenkään selvää, että tieteellisen tutkimuksen päämääränä olisi havainnollistaa näitä osia sinänsä. Krickel kuitenkin näyttäisi tiedostavan tämän haasteellisuuden (2017, 467; 2018, 65-66), mutta selkeää vastausta ongelmaan ei näytä ainakaan toistaiseksi löytyvän. Krickel näyttäisi pikemminkin ohittavan haasteen toteamalla antaneessa vain metafyyssisen analyysin tasojen välisestä kausaalisuudesta. Toisaalta MM-ehdon pelastava puhdas metafyyssinen analyysi, jota ei motivoi sovellettavuus tieteeseen, saattaa osoittautua tieteenfilosofisesta näkökulmasta käsin varsin hedelmättömäksi. Oleellista on lisäksi huomata, että Krickelin ilmiötermi on tapahtuman alainen entiteetti (EIO) sinänsä, mikä on varsin omituinen käsitevalinta tieteellisen käytännön kannalta. Väitän kuitenkin, ettei ole mikään selviö, että Krickelin metafyyssiset käsitteet vastaisivat niitä oletuksia, joita tieteilijät tekevät ilmiöistä. Väittäisin, että tutkijat kohtelevat ilmiöitä enemmänkin seikkoina, jotka ilmentävät kausaalista kapasiteettiaan (disposiota), koska niillä on tietynlainen konstitutionaalinen rakenne (ks. tämän tarkastelun osio 2.3.2). Jos olen oikeassa, silloin ei ole välttämättä mielekästä puhua ilmiöstä tapahtuman alaisena entiteettinä. Viimeiseksi todettakoon, että Krickelin kannan suhteen haasteellisia ovat tapaukset, joissa komponentin ja ilmiön muutokset ilmenevät samanaikaisesti. Toisin sanoen tämä on käänteinen ongelma kuin abduktiivisessa mallissa, jossa ongelmana oli se, ettei siinä voida käsitteellistää muutoksia, jotka eivät tapahdu samanaikaisesti makro ja mikrotasoilla. Tämä on kuitenkin jotakin, mitä Krickelin kanta ei voi ottaa huomioon.

6.3. Mutaalisen manipulaation tulevaisuudennäkymät

Tarkastelun pohjalta voidaan spekuloida MM-ehdon asemaa konstitutionaalisena relevanssiehtona. Tässä on esitettyä kritiikkiä abduktiivisen että kausaalisen ratkaisuyrityksien suhteen ei voida pitää millään tavalla ratkaisevana. Pikemminkin sitä voidaan pitää filosofisena evidenssinä sen suhteen, että tarjotut ratkaisuyritykset ovat nykyisissä muodoissaan ainakin puutteellisia. Yhtäältä voidaan nimittäin huomata, että molemmat ratkaisuehdotukset ovat hedelmällisiä sikäli kuin ne avaavat uusia kehityssuuntia MM-ehdon analysoimisen suhteen. Varsinkin Krickelin ansiona voidaan

pitää metafysiikan mukaan tuomista keskusteluun. Hän on onnistuneesti useissa yhteyksissä (Kaiser & Krickel 2017; Krickel 2017; 2018) tuonut hyvin ilmi, kuinka uusi mekanistisen filosofia tarvitsee metafysiistä ”hienosäätöä”.

Toisaalta voidaan tehdä kuitenkin se kriittinen huomio, että MM-ehdon ratkaisuehdotuksia voidaan pitää hyvin raskaina. Kumpikin käsitellyistä ehdotuksista liikkuvat hyvin abstraktilla tasolla. Näin on laita varsinkin Krickelin ratkaisun suhteen, jossa joudutaan määrittelemään ilmiön käsite täysin uudella tavalla, joka on varsin vieras tieteenfilosofialle. On selvää, että filosofiassa on pitkälti kyse määritelmien antamisesta, niiden kehittämisestä ja analysoimisesta. Tieteenfilosofia ei ole tässä poikkeus. Mikäli tieteenfilosofian halutaan edustavan tiedettä sellaisena kuin se on, herättää tämän kysymyksen siitä, millainen määritelmien antamisen vapaus tieteenfilosofilla on. Väitän, että Krickel ottaa liian suuren vapauden määritellessään MM-ehdon ikään kuin ”metafysiikka edellä”. Tämän voidaan katsoa olevan kyseenalainen lähestymistapa sikäli kuin se voi lipua tilanteeseen, jossa jokin tieteen osa määritellään lähestulkoon *ad hoc*-tyyppisesti sellaiseksi, että se tyydyttää vain filosofiset päämäärät. Väitän siis, että mikäli MM-ehto ei ole pelastettavissa muuta kuin *ad hoc* tai vetoamalla epäkäytännöllisiin johtopäätöksiin, silloin hedelmällisempää on siirtyä tulevaisuudessa tarkastelemaan vaihtoehtoisia tapoja konstitutionaalisen riippuvaisuuden käsittelemiseksi. Yksi tuoreimmista vaihtoehtoista lienee konstitutionaalisen relevanssin määrittelemisen käyttäen hyväksi Bayesin verkkoanalyysia ja Markov faktorisointia (*Markov factorization*), mikä ei ole sitoutunut mutuaaliseen manipulaatioon sinänsä (Gebharder 2016; 2017, 115-175). Tähän analyysiin ei voida kuitenkaan enää tarttua tässä yhteydessä. Jätettäköön se tulevan tutkimuksen aiheeksi.

7. Päätelmiä

Miten tarkastelussa ollaan onnistuttu vastaamaan sille asetettuihin haasteisiin H1-H3? Ensimmäinen näistä haasteista (H1) koski sitä, mikä selittää maailman olemisen niin- tai-näin. Olen esittänyt osiossa 2 ja 3 ontisen selitysnäkemyksen mukaisesti, että maailman kausaaliset rakenteet eli mekanismit selittävät ilmiöt. Määrittelin mekanismit tapahtuman alaisten entiteettien järjestäytyneiksi, ei-aggregatiivisia kokonaisuuksiksi, jotka sellaisenaan vastaavat ilmiöistä. En voi väittää tämän olevan tyhjentävä määritelmä. Katsoisin, että tulevan tutkimuksen kannalta aiheellista on keskittyä varsinkin määritelmän metafysiseseen analyysiin. Se, mitä tarkoitetaan, tapahtumilla,

entiteeteillä, ilmiöstä vastaamisella ja ilmiöllä itsellään mekanistisessa analyysissä, ovat jotakin, mikä kaipa vielä selkoperäisempää tarkastelua. Toki tämä on haasteellista. Olen jo tässä tarkastelussa osoittanut varsinkin Krickelin kohdalla, kuinka haastavia metafyyysiset uudelleen määrittelyt voivat olla tieteenfilosofiassa. Väittäisin, että käsitteiden filosofinen uudelleen määrittely on usein altis vaikutukselle, joka etäännyttää filosofin ja hänen määritelmänsä tieteellisestä käytännöstä. En väitä tämän olevan uusi ongelma, mutta väitän sen olevan yhä ajankohtainen ongelma tieteenfilosofiassa.

Entä kuinka on laita toisen haasteen (H2) suhteen? Tarkemmin sanoen millä keinoilla ja missä muodossa tieteellinen selitys voidaan ja tulisi muotoilla? Tartuin tarkastelussa haasteeseen esittelemällä, kuinka mekanistiset mallit ovat keskeisiä representaatiovälikappaleita selityksellisen tekstin, maailman (mekanismin) ja selittävän agentin välillä. Mekanistinen malli ei ole siis selitys sinänsä; mekanismi on se, *mikä* selittää ja mekanistinen malli on se, *millä* selitetään. Osoitin kuitenkin osiossa 4, kuinka ontisuus on keskeinen mallintamisen kannalta, sillä ilman ontisuutta putoaa mallilta normatiivinen pohja. Tämän johdosta esitin väitteen, että malli on *selityksellinen malli*, jos sen sisäisten muuttujien väliset riippuvaisuussuhteet representoivat maailman kausaalisia riippuvaisuussuhteita. Tähän vedoten havainnollistin osiossa 4, vastoin dynaamisia sovinisteja, kuinka mallin selitysvaikutus ei nouse mallin sisäisten muuttujien välisistä matemaattista riippuvaisuuksista. Haluan korostaa sitä, että en väitä, etteikö matemaattisten riippuvaisuussuhteiden jäljittämiseksi olisi mitään tieteellisesti mielenkiintoista. Päinvastoin, uskon tämän olevan erittäin hedelmällinen mallintamisen tapa. Voidaan kysyä, miksi matemaattiset riippuvaisuussuhteet pätevät niin hyvin. Hedelmällisen tästä kysymyksestä tekee se, että ne pätevät todennäköisesti siksi, koska jotakin kausaalista on tekeillä. Malli ei kuitenkaan selitä matemaattisten riippuvaisuussuhteiden nojalla; se selittää, koska osa näistä riippuvaisuussuhteista representoi mallin ulkoisia, todellisuuden kausaalisia riippuvaisuussuhteita. En kuitenkaan väitä ratkaisseeni haastetta H2. Varsinkin abstrahoinin ja idealisoinnin tarkempi yhteensovittaminen ontisuuden kanssa on jotakin, mihin tulevaisuuden tutkimuksen tulisi kiinnittää huomiota. En väitä *todistaneeni* tätä yhteensovittuvuutta tässä, mutta väittäisin osoittaneeni, että yhteensovittaminen on mahdollista ja välttämätöntä.

Kolmas haaste (H3) koski sitä, mikä tekee jostakin asiasta selityksellisesti keskeisen eli relevantin. Osoitin, että tähän haasteeseen voidaan tarttua osoittamalla se, mikä riippuu

(kausaalisesti) mistäkin. Osiossa 5 osoitin, että Woodwardin interventionismi tarjoaa hedelmällisen tavan käsitteellistää nämä riippuvaisuussuhteet kausaalisen selittämisen kannalta; selityksellisesti relevantit riippuvaisuussuhteet ovat kausaalisia eron tekeviä suhteita, joiden välinen eron tekevä suhde pysyy invarianttina joidenkin aktiivisten kontrafaktuaalien (interventioiden) suhteen. Esitin, että relevantit seikat ovat näin ollen sellaisia, jotka auttavat muodostamaan *mitä-jos*-päätelmiä. Osoitin osiossa 6, että tämän käsityksen soveltaminen konstitutionaaliseen selittämiseen on hyvin haasteellista MM-ehdon muodossa, sillä näyttäisi siltä, ettei konstitutionaalisten tasojen välillä voi vallita symmetristä kausaalista eron tekevää suhdetta. Osoitin, että abduktiivinen ja kausaalinen lähestymistapa eivät kykene ongelmitta ratkaisemaan tätä MM-ehdon symmetrisuushaastetta.

Tarkastelun perusteella uskallan jopa väittää, että H3-haaste on konstitutionaalisten selittämisen kannalta uuden mekanistisen filosofian vakavampia ja ehkä keskeisempiä haasteita. Lienee selvää tarkastelun perusteella, että MM-ehto ei ole välttämättä paras mahdollinen tapa lähestyä tätä haastetta tulevaisuudessa. Toki on syytä muistaa, ettei MM-ehto ole ainoa mahdollinen lähestymistapa. Tässä tarkastelussa ei valitettavasti ole esitetty vaihtoehtoisia näkökulmia, mutta uskon, että vaihtoehtoisten näkökulmien kehittäminen tulee tarjoamaan MM-ehdotta hedelmällisempiä ja kenties käytännöllisempiä lähestymistapoja haasteen ratkaisemiseksi. Toistaiseksi voidaan kuitenkin todeta, ettei mekanistien keskuudessa ole selvää käsitystä siitä, minkälaisella teorialla H3-haastetta tulisi lähestyä konstitutionaalisen relevanssin suhteen.

8. Lähdeluettelo

Baumgartner, M. & Casini, L., 2017. An Abductive Theory of Constitution. *Philosophy of science*, 84(2), 214-233.

Baumgartner, M. & Gebharder, A., 2016. Constitutive Relevance, Mutual Manipulability, and Fat-Handedness. *The British Journal for the Philosophy of Science*, 67(3), 731-756.

Baumgartner, M. & Wilutzky, W., 2017. Is it possible to experimentally determine the extension of cognition?. *Philosophical Psychology*, 30(8), 1104-1125.

Bechtel, W., 2008. *Mental Mechanisms*. New York: Routledge.

- Bechtel, W. & Abrahamsen, A., 2005. Explanation: a mechanist alternative. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421-441.
- Bechtel, W. & Levy, A., 2013. Abstraction and the Organization of Mechanisms. *Philosophy of Science*, 80(2), 241-261.
- Bogen, J., 2005. Regularities and causality; generalizations and causal explanations. *Studies in History and Philosophy of Biology & Biomed Science*, 36(2), 397-420.
- Bromberger, S., 1966. Why-Questions. Teoksessa: *Mind and Cosmos*. Pittsburgh: University of Pittsburgh Press, 86-111.
- Bromberger, S., 1992. *On What We Know*. Chicago: The University of Chicago Press.
- Cartwright, N., 1983. *The Truth Doesn't Explain*. Clarendon: Oxford University Press.
- Chalmers, D., 2008. Strong and Weak Emergence. Teoksessa: P. Clayton & P. Davies, toim. *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*. Oxford: Oxford University Press, 245-254.
- Chemero, A. & Silberstein, M., 2008. After the Philosophy of Mind: Replacing Scholasticism with Science. *Philosophy of Science*, 75(1), 1-27.
- Chirimuuta, M., 2014. Minimal models and canonical neural computations: the distinctness of computational explanation in neuroscience. *Synthese*, 191(2), 127-153.
- Coffa, A., 1974. Hempel's Ambiguity. *Synthese*, 28(2), 141-163.
- Craver, C., 2006. When mechanistic models explain. *Synthese*, 153(3), 355-376.
- Craver, C., 2007. *Explaining The Brain*. Oxford: Oxford University Press.
- Craver, C., 2014. The Ontic Account of Scientific Explanation. Teoksessa: M. Kaiser, R. Scholz & D. Plenge, toim. *Explanation in the Special Sciences*. London: Springer, 27-55.
- Craver, C. & Bechtel, W., 2007. Top-down causation without top-down causes. *Biology and Philosophy*, 22(4), 547-563.
- Craver, C. & Darden, L., 2013. *In Search of Mechanisms*. Chicago: The University of Chicago Press.

- Craver, C. & Lindley, D., 2001. Discovering Mechanisms in Neurobiology: The Case of Spatial Memory. Teoksessa: P. Machamer, R. Grush & P. McLaughlin, toim. *Theory and Method in the Neurosciences*. Pittsburgh: University of Pittsburgh, 112-137.
- Darden, L., 2006. *Reasoning in biological discoveries*. Cambridge : Cambridge University Press.
- Darden, L., 2008. Thinking Again about Biological Mechanisms. *Philosophy of Science*, 75(5), 958-969.
- Darden, L. & Craver, C., 2002. Strategies in the interfield discovery of the mechanism of protein synthesis. *Studies in History and Philosophy of Biology & Biomedical Science*, 33(1), 1-28.
- Gebharder, A., 2017. *Causal Nets, Interventionism, and Mechanisms*. Cham: Springer.
- Gebharder, A., 2017. Uncovering constitutive relevance relations in mechanisms. *Philosophical Studies*, 174(11), 2645-2666.
- Giere, P., 2004. How models are used to represent reality. *Philosophy of Science*, 71(5), 742-752.
- Giere, R., 1999. Using Models to Represent Reality. Teoksessa: L. Magnani, N. Nersessian & P. Thagard, toim. *Model-Based Reasoning in Scientific Discovery*. New York: Springer, 41-57.
- Glennan, S., 1996. Mechanism and the Nature of Causation. *Erkenntnis*, 44(1), 49-71.
- Glennan, S., 2002. Rethinking Mechanistic Explanation. *Philosophy of Science*, 69(3), 342-353.
- Glennan, S., 2005. Modeling mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 443-464.
- Glennan, S., 2017. *The New Mechanical Philosophy*. Oxford: Oxford University Press.
- Hempel, C., 1965. *Aspects of Scientific Explanation*. New York: Free Press.
- Hempel, C. & Oppenheim, P., 1948. Studies in the Logic of Explanation. *Philosophy of Science*, 15(2), 135-175.

- Illari, P., 2013. Mechanistic Explanation: Integrating the Ontic and Epistemic. *Erkenntnis*, 78(2), 237-255.
- Illari, P. & Williamson, J., 2012. What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2(1), 119-135.
- Jones, M., 2005. Idealization and Abstraction: A Framework. Teoksessa: M. Jones & N. Cartwright, toim. *Idealization XII : correcting the model : idealization and abstraction in the sciences*. New York: Rodopi, 173-219.
- Kaiser, M. & Krickel, B., 2017. The Metaphysics of Constitutive Mechanistic Phenomena. *The British Journal for the Philosophy of Science*, 68(3), 745-779.
- Kaplan, M., 2011. Explanation and description in computational neuroscience. *Synthese*, 183(3), 339-373.
- Kaplan, M. & Craver, C., 2011. The Explanatory Force of Dynamical and Mathematical Models in Neuroscience: A Mechanistic Perspective.. *Philosophy of science*, 78(4), 601-627.
- Koskinen, R., 2017. Synthetic biology and the search for alternative genetic systems: Taking how-possibly models seriously. *European Journal for Philosophy of Science*, 7(3), 493-506.
- Krickel, B., 2017. Making Sense of Interlevel Causation in Mechanisms from a Metaphysical Perspective. *Journal for General Philosophy of Science*, 48(3), 453-468.
- Krickel, B., 2018. Saving the mutual manipulability account of constitutive relevance. *Studies in History and Philosophy of Science*, 68, 58-67.
- Kuorikoski, J. & Ylikoski, P., 2015. External representations and scientific understanding. *Synthese*, 192(12), 3817-3837.
- Leuridan, B., 2012. Three Problems for the Mutual Manipulability Account of Constitutive Relevance in Mechanisms. *The British Journal for the Philosophy of Science*, 63(3), 399-427.
- Lewis, D., 1986. *Philosophical Papers 2*. Oxford: Oxford University Press.
- Levy, A. & Bechtel, W., 2013. Abstraction and the Organization of Mechanisms. *Philosophy of Science*, 80(2), 241-261.

- Love, A. & Nathan, M., 2015. The Idealization of Causation in Mechanistic Explanation. *Philosophy of Science* , 82(5), 761-774.
- Machamer, P., 2004. Activities and Causation: The Metaphysics and Epistemology of Mechanisms. *International Studies in the Philosophy of Science*, 18(1), 27-39.
- Machamer, P., Darden, L. & Craver, C., 2000. Thinking about Mechanisms. *Philosophy of Science*, 67(1), 1-25.
- Mitchell, S., 1997. Pragmatic Laws. *Philosophy of Science*, 64, 468-479.
- Mitchell, S., 2000. Dimensions of Scientific Law. *Philosophy of Science*, Osa/vuosikerta 67, 242-265.
- Mäki, U., 2005. Models are experiments, experiments are models. *Journal of Economic Methodology*, 12(2), 303-315.
- Mäki, U., 2009. MISSing the World. Models as Isolations and Credible Surrogate Systems. *Erkenntnis*, 70(1), 29-43.
- Piccinini, G., 2006. Computational explanation in neuroscience. *Synthese*, 153(3), 343-353.
- Raerinne, J., 2011. *Generalizations and models in ecology : lawlikeness, invariance, stability, and robustness*. Helsinki: Helsingin yliopisto, filosofian, historian, kulttuurin ja taiteiden tutkimuksen laitos. Väitöskirja.
- Raerinne, J., 2018. Abstraction in ecology: reductionism and holism as complementary heuristics. *European Journal for Philosophy of Science*, 8(3), 395-416.
- Romero, F., 2015. Why there Isn't inter-level causation in mechanisms. *Synthese*, 192(11), 3731-3755.
- Salmon, W., 1971. Statistical Explanation. Teoksessa: W. Salmon, J. Greeno & R. Jeffrey, toim. *Statistical explanation & statistical relevance*. Pittsburgh: University of Pittsburgh Press, 29-88.
- Salmon, W., 1984. *Scientific Explanation and the Causal Structure of the World..* Princeton: Princeton University Press.
- Salmon, W., 1989. *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press.

- Salmon, W., 1998. *Causality and explanation*. New York: Oxford University Press.
- Scheines, R., 2005. The Similarity of Causal Inference in Experimental and Non-experimental Studies. *Philosophy of Science* , 75(5), 927-940.
- Scriven, M., 1959. Explanation and Prediction in Evolutionary Theory. *Science*, 130(3374), 477-482.
- Thagard, P., 1998. Explaining Disease: Causes, Correlations, and Mechanisms. *Minds and Machines*, 8(1), 61–78.
- Ward, L., 2002. *Dynamical cognitive science*. Cambridge: MIT Press.
- Waskan, J., 2011. Mechanistic Explanation at the Limit. *Synthese*, 183(3), 389-408.
- Woodward, J., 2000. Explanation and invariance in the special sciences. *The British Journal for the Philosophy of Science*, 51(2), 197-254.
- Woodward, J., 2001. Law and Explanation in Biology: Invariance Is the Kind of Stability that Matters. *Philosophy of Science*, 68(1), 1–20.
- Woodward, J., 2002. What Is a Mechanism? A Counterfactual Account. *Philosophy of Science*, 69(3), 366-377.
- Woodward, J., 2003. *Making things happen: a theory of causal explanation*.. Oxford: Oxford University Press.
- Woodward, J., 2011. Mechanisms revisited. *Synthese*, 183(3), 409-427.
- Woodward, J., 2015. Interventionism and causal exclusion. *Philosophy and Phenomenological Research*, 91(2), 303-347.
- Woodward, J. & Hitchcock, C., 2003. Explanatory Generalizations, Part I: A Counterfactual Account. *Noûs*, 37(1), 1-24.
- Wright, C., 2012. Mechanistic explanation without the ontic conception. *European Journal for Philosophy of Science*, 2(3), 375-394.
- Ylikoski, P., 2013. Causal and Constitutive Explanation Compared. *Erkenntnis*, 78(2), 277-297.
- Ylikoski, P. & Kuorikoski, J., 2010. Dissecting explanatory power. *Philosophical Studies*, 148(2), 201-219.

